CHARACTERIZING THE FUNCTION OF CLOCKWORK ORANGE IN THE CIRCADIAN FEEDBACK LOOPS IN *DROSOPHILA MELANOGASTER*

A Dissertation

by

JIAN ZHOU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Paul E. Hardin
Committee Members,	Deborah Bell-Pedersen
	Luis René García
	Vladislav M. Panin
Head of Department,	Thomas D. McKnight

August 2017

Major Subject: Biology

ABSTRACT

The *Drosophila* circadian oscillator controls daily rhythms in physiology, metabolism and behavior via transcriptional feedback loops. CLOCK-CYCLE (CLK-CYC) heterodimers initiate feedback loop function by binding enhancer box (E-box) elements to activate *period* (*per*) and *timeless* (*tim*) transcription. PER-TIM heterodimers then accumulate, bind CLK-CYC to inhibit transcription, and are ultimately degraded to enable the next round of transcription. Although tremendous efforts have been made to understand how these feedback loops are regulated, the detailed molecular mechanism of transcriptional repression is still not clear. By using genetic, molecular, and genome-wide bioinformatic analyses, I have characterized the molecular function of the transcription factor CLOCKWORK ORANGE (CWO) for the circadian transcriptional repression, and further identified its downstream targets aiming to understand its function at genome-wide level.

The timing of transcriptional events in the circadian feedback loops coincide with, and are controlled by, rhythms in CLK-CYC binding to E-boxes. PER rhythmically binds CLK-CYC to initiate transcriptional repression, and subsequently promotes the removal of CLK-CYC from E-boxes. However, little is known about the mechanism by which CLK-CYC are removed from DNA. Here I show that the transcription repressor CWO rhythmically binds E-boxes upstream of core clock genes in a reciprocal manner to CLK, thereby promoting PER-dependent removal of CLK-CYC from E-boxes, and maintaining repression until PER is degraded and CLK-CYC

displaces CWO from E-boxes to initiate transcription. These results suggest a model in which CWO co-represses CLK-CYC transcriptional activity in conjunction with PER by competing for E-box binding once CLK-CYC-PER complexes have formed. Given that CWO orthologs DEC1 and DEC2 also target E-boxes bound by CLOCK-BMAL1, a similar mechanism may operate in the mammalian clock. Furthermore, using ChIP-seq and RNA-seq analyses, I show that CWO directly and indirectly regulates gene expression at genome-wide level. A substantial overlap between CWO and CLK direct target genes suggests that CWO plays a potential role in regulating CLK-mediated transcription globally. Moreover, CWO indirectly regulates a subset of genes encoding kinases and phosphatases at transcriptional level, suggesting its role in the posttranscriptional regulation of CLK during the circadian cycle.

DEDICATION

This work is dedicated to Charles Darwin, Stephen Jay Gould and Richard Feynman, whose books lured me towards science; and to Gustav Mahler, Anton Bruckner, Stanley Kubrick, Fyodor Dostoyevsky, Johann Wolfgang von Goethe and many others who saved me from becoming a scientific nerd.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Paul Hardin, for his guidance and supports. He is a nice gentleman, a great mentor and an open mind scientist. He always encourages me to try new things and to use the newly developed tools and techniques to solve the tough problems in my projects. It is a great pleasure to work with him in his lab that has a perfect balance in traditional molecular genetics and the most cutting-edge biological techniques.

I would also like to thank my committee members: Dr. Deborah Bell-Pedersen, Dr. Luis René García and Dr. Vlad M. Panin, for their guidance and supports throughout the course of this research. I want to specially thank Dr. Wangjie Yu, a former research assistant professor in the lab. He has been a great scientific consultant, not only helped me with my project, but also taught me how to think logically and creatively. In the lab, he was almost my second mentor, and outside the lab he is also a good friend to me. My project would be impossible without his help.

I would like to thank two young professors, Dr. Jerome Menet and Dr. Christine Merlin, a lovely couple in our department. They generously shared their experience of genome-wide research with me and taught me all the bioinformatic knowledge and skills that are not only helpful for me project, but also valuable for my career development in the future.

I am grateful to the previous and current members of our lab, Dr. Guruswamy Mahesh, Dr. Jerry Houl, Dr. Wangjie Yu, Paul Kim, Kushan Gunawardhana, Courtney

Caster, Dr. Parul Agrawal, Tianxin Liu and Stephanie Durkacz for training and supporting me through these years of my graduate study. Thanks also go to my friends Shian Liu, Yi Sun, Yumeng Fan, Muwei Zhang, Xin Chen, Wangming Ji and many others for their help and encouragement.

Finally, I would like to thank my mother and father for their caring and support. Both of them knew nothing about biology at the beginning and now after these years I feel that they know something about the circadian clock already. In the end, I must say that it is my wife's unique encouragement: "You experiment is not going to work…again…" that made me went through all these years of research and finally made it work!

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Paul E. Hardin and Professors Deborah Bell-Pedersen and Luis René García of the Department of Biology and Professor Vlad M. Panin of the Department of Biochemistry and Biophysics.

All work for the dissertation was completed independently by the student, under the advisement of Dr. Paul Hardin of the Department of Biology.

Funding Sources

There are no outside funding contributions to acknowledge related to the research and compilation of this document.

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

CONTRIBUTORS AND FUNDING SOURCES vii

TABLE OF CONTENTS viii

LIST OF FIGURES x

LIST OF TABLES xii

CHAPTER I INTRODUCTION 1

I.A. Biological clocks are ubiquitous and beneficial in multiple organisms 1
I.B. The molecular basis of the circadian clock 2
I.C. The *Drosophila* circadian molecular oscillator 5
I.C.1. The core feedback loop 5
I.C.2. Interlocked feedback loops 6
I.C.3. Posttranscriptional regulation of circadian transcription in *Drosophila* 8
I.D. Transcription repression in circadian feedback loops in *Drosophila* 11
I.E. CWO is a circadian transcriptional regulator in *Drosophila* 13
I.E.1. *cwo* transcription is controlled by CLK through canonical E-boxes 13
I.E.2. *cwo* encodes a transcriptional factor that specifically binds E-boxes *in vitro* 15
I.E.3. *cwo* is a core clock component and negatively regulates the expression of clock genes 17
I.F. The goal of the study 20

CHAPTER II MATERIALS AND METHODS 22

II.A. Transgene construction and transgenic fly generation 22
II.B. ChIP Protocol 24
II.C. Quantitative RT-PCR 36
II.D. Western blotting 37
II.E. ChIP-seq and RNA-seq libraries preparation 38
II.E.1. ChIP-seq library preparation 38
II.E.2. RNA-seq library preparation 38
II.F. Bioinformatic analyses for ChIP-seq and RNA-seq data 39

II.F.1. ChIP-seq mapping and peak finding 39
II.F.2. RNA-seq data analysis 40
II.G. *Drosophila* activity monitoring and behavior analysis 40

CHAPTER III CWO ENHANCES PER MEDIATED RHYTHMS IN TRANSCRIPTIONAL REPRESSION BY COMPETING WITH CLK IN E-BOX BINDING 41

III.A. Introduction 41
III.B. Results 43
III.B.1. CWO is present at constant levels and rhythmically binds E-boxes in a reciprocal pattern compared to CLK 43
III.B.2. CWO and CLK bind tandem E-boxes upstream of *tim* 47
III.B.3. CWO represses CLK binding to *tim* promoter during transcription repression 51
III.B.4. PER is required for CWO to displace CLK-CYC binding on E-boxes 53
III.C. Summary 57

CHAPTER IV GENOME-WIDE ANALYSIS OF CWO'S FUNCTION IN TRANSCRIPTIONAL REGULATION 58

IV.A. Introduction 58
IV.B. Results 60
IV.B.1. HA tagged CWO protein is detectable by anti-HA and can partially rescue *cwo* mutant phenotype 60
IV.B.2. Identification of CWO direct target genes in *Drosophila* 62
IV.B.3. Differential gene expression in wild-type vs. cwo^{5073} mutant flies 71
IV.B.4. CWO potentially regulates CLK phosphorylation to promote transcription 74
IV.C. Summary 78

CHAPTER V DISCUSSION AND SUMMARY 80

V.A. CWO competes with CLK in E-box binding to enhance PER mediated rhythms in transcriptional repression 80
V.B. Genome-wide analysis of *cwo* function in transcriptional regulation 86
V.C. Conclusions 91

REFERENCES 93

APPENDIX 110

LIST OF FIGURES

Page

Fig I.1. Comparison of mammalian and *Drosophila* oscillators (taken with permission from Hardin, 2004) 4

Fig I.2. Two transcription feedback loops of the *Drosophila* Clock (Adapted with permission from Cyran et al., 2003) 7

Fig I.3. Posttranscriptional regulation of the Per/Tim feedback loop (Taken with permission from Hardin, 2005) 10

Fig I.4. PER dependent rhythmic binding of CLK on E-boxes (Adapted with permission from Yu et al., 2006) 12

Fig I.5. *cwo* transcription is rhythmically activated by CLK (Adapted with permission from Lim et al., 2007; Matsumoto et al., 2007). 14

Fig I.6. CWO protein directly targets known clock genes (Taken with permission from Matsumoto et al., 2007). 16

Fig I.7. CWO represses the expression of clock genes in cell culture and *in vivo* (Taken with permission from Kadener et al., 2007). 19

Fig III.1. CWO protein is present at constant levels in fly heads. 45

Fig III.2. CWO rhythmically binds E-boxes of core clock genes in antiphase to CLK... 46

Fig III.3. CWO and CLK bind the same tandem E-boxes in the *tim* circadian enhancer. 48

Fig III.S1. *tim* promoter fragments bearing E-box mutations abolish mRNA cycling. ... 50

Fig III.4. CWO reduces CLK binding to *tim* E-boxes during transcriptional repression. 52

Fig III.5. PER is required for CWO to compete with CLK-CYC for E-box binding. 54

Fig III.6. Model for PER dependent binding competition between CWO and CLK-CYC on E-boxes. 56

Fig IV.1. CWO-HA protein is present at constant levels in *cwo-HA; cwo*^{5073} transgenic fly heads. 61

Fig IV.2. ChIP-seq analysis of CWO binding sites and comparison to previous microarray data. 64

Fig IV.3. Visualizing of CLK and CWO binding to the core clock genes. 65

Fig IV.4. Motif enriched in CWO binding regions 66

Fig IV.5. ChIP-seq analysis of CLK binding sites and comparison to CWO ChIP-seq and previous microarray data. 71

Fig IV.6. RNA-seq in WT vs. *cwo*^{5073} mutant identified differentially expressed mRNAs. 73

Fig IV.7. Phosphorylation of PER and CLK protein in wild-type and *cwo*^{5073} transgenic flies at different time points of the circadian cycles. 77

Fig IV.8. CWO indirectly regulates the mRNA levels of *CG17746*, *trbl* and *Hex-C*. 78

LIST OF TABLES

Page

Table IV.1. *cwo*-HA partially rescues *cwo*^{5073} mutant phenotype. 61

Table.IV.2. Overall view of the CWO and CLK ChIP-seq peaks in wild-type fly heads. 63

Table IV.3. GO (biological process) analysis and display of CWO binding peaks. 68

Table IV.4 GO (functional) analysis and display of CWO binding peaks. 69

Table IV.3. Kinases and phosphatases that are upregulated or downregulated in *cwo*^{5073} mutant. 74

CHAPTER I

INTRODUCTION

I.A. Biological clocks are ubiquitous and beneficial in multiple organisms

We live on a planet that has diurnal cycles governed by the earth's rotation once in about 24 hours. In order to adapt to this light-dark environment, almost all organisms from Cyanobacteria to humans harbor biological clocks that allow them to anticipate external environmental changes (Bell-Pedersen et al., 2005). This biological timekeeping system is not passively driven by environmental cycles, but controlled by endogenous circadian clocks that keep time in the absence of environmental cues. Importantly, the clock is known to control numerous physiological and molecular processes, as there is growing evidence that clock deficiencies are associated with abnormal sleep-wake cycles (e.g., Familial Advanced Sleep Phase Syndrome), metabolic diseases and mood disorders (Chaix et al., 2014; Chung et al., 2014; He et al., 2009; Patke et al., 2017; Xu et al., 2007). A more recent study indicated that the majority of best-selling drugs and World Health Organization essential medicines directly target the products of rhythmically expressed genes, emphasizing the clinical importance of understanding the molecular organization of the circadian system (Zhang et al., 2014).

The master pacemaker of circadian clock in mammals and humans is the suprachiasmatic nuclei, a set of paired structures containing approximately 10,000 neurons each, and are located on either side of the third ventricle in the anterior hypothalamus (Moore and Eichler, 1972; Stephan and Zucker, 1972). In *Drosophila*, the

circadian clock operates in ~75 pacemaker neurons per hemisphere that function to drive activity rhythms (Helfrich-Förster, 2003). At the molecular level, both *Drosophila* and mammalian circadian clocks are composed of highly conserved cell-autonomous transcriptional feedback loops that contain positive and negative elements to regulate cyclical gene expression (reviewed in Dunlap, 1999). This conserved timekeeping system among different species thus allows us to investigate the molecular clocks in complex human brains by studying more approachable model organisms, such as *Drosophila*. Therefore, the powerful molecular and genetic tools that have been utilized for decades in *Drosophila* made it a good system in chronobiological research for more than forty years.

I.B. The molecular basis of the circadian clock

In eukaryotes, the circadian clock keeps time via one or more transcriptional feedback loops (Bell-Pedersen et al., 2005). The molecular oscillators of *Drosophila* and mammals both contain feedback loops that drive rhythmic transcription in the opposite phases of the circadian cycle (Young and Kay, 2001), and share many features and conserved components (Fig I.1). In *Drosophila*, a heterodimer formed by basic-helix-loop-helix (bHLH) PER-ARNT-SIM (PAS) partners CLOCK (CLK) and CYCLE (CYC) binds E-box enhancer elements to activate transcription of target genes via two feedback loops. In the first "core timekeeping" or "core" loop, the CLK/CYC target genes are *period* (*per*) and *timeless* (*tim*) (Allada et al., 1998; Darlington et al., 1998; Hao et al., 1997; Rutila et al., 1998), and in the second "interlocked" loop the target

genes are *vrille* (*vri*) and *Par Domain Protein 1ε (Pdp1ε)* (Cyran et al., 2003; Glossop et al., 2003). These positive and negative factors are largely conserved in mammals: CLK ortholog, CIRCADIAN LOCOMOTOR OUTPUT CYCLES KAPUT (CLOCK), forms a heterodimer with the CYC ortholog BRAIN and MUSCLE ARNT-LIKE PROTEIN 1 (BMAL1) to form positive factors, whereas PER orthologs, PERIOD (PER1and PER2), form a heterodimer with CRYPTOCHROME (CRY1 and CRY2) rather than TIM to form the negative factor (Lowrey and Takahashi, 2011). Both the CLK/CYC heterodimer and the CLOCK/BMAL1 heterodimer are at the core of the oscillator for *Drosophila* and mammals, respectively (Darlington et al., 1998; Gekakis et al., 1998; Hogenesch et al., 2000; Rutila et al., 1998). In the core loop the CLOCK/BMAL1 target genes are *Per (Per1 & Per2)* and *Cry* (*Cry1 & Cry2)*, in the second loop their target genes are *Rev-erbs* and *retinoid-related orphan receptors* (*Rors*). The first feedback loop is similar in *Drosophila* and mammals, with the PER/TIM complex repressing CLK/CYC transcription in *Drosophila* (Bae et al., 1998; Chang and Reppert, 2003) and the PER/CRY complex repressing CLOCK/BMAL1 transcription in mammals (Kume et al., 1999). In the second loop, PDPlε activates *Clk* and VRI represses *Clk* in *Drosophila*, which is also similar to the case in mammals, as RORα activates *Bmal1* and REV- ERBα represses *Bmal1* (Cyran et al., 2003; Glossop et al., 2003; Preitner et al., 2002; Sato et al., 2004; Ueda et al., 2002) (Fig I.1). Given these conserved features between *Drosophila* and mammals, gaining insight into the *Drosophila* circadian clock will shed light on understanding the oscillator of mammals and humans, which could lead to beneficial treatments of circadian disorders and other related diseases.

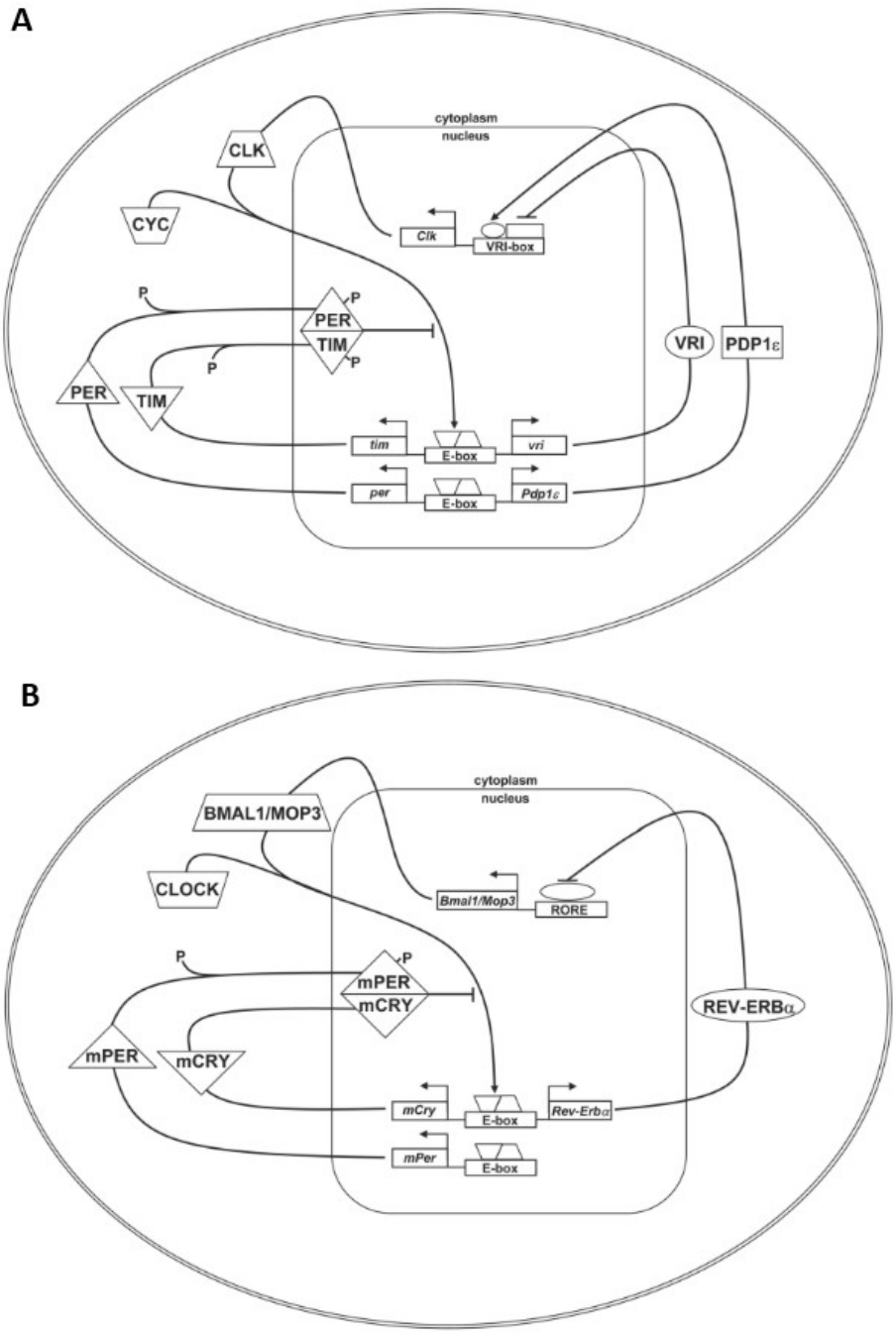


Fig I.1. Comparison of mammalian and *Drosophila* oscillators (taken with permission from Hardin, 2004)

The molecular framework of **(A)** *Drosophila* circadian clock and **(B)** mammalian circadian clock.

I.C. The *Drosophila* circadian molecular oscillator

Genetic study of circadian behaviors of *Drosophila*, primarily rhythmic locomotor activity, has uncovered the components of the molecular clock, which form two transcriptional feedback loops and are vital for biological timekeeping. Here I will briefly introduce the molecular mechanism of these two feedback loops.

I.C.1. The core feedback loop

In the core loop, also termed the *per/tim* loop, CLK-CYC activate *per* and *tim* transcription during mid-day starting at ~ZT4 (Zeitgeber Time, or ZT, refers to time in hours during a light-dark cycle, where ZT0 is lights on and ZT12 is lights off), which effects a rise in *per* and *tim* mRNA levels that peak during the early evening. The 69 base pair (bp) circadian regulatory sequence (CRS) of *per* that contains E-box is sufficient to drive circadian oscillations of transcription *in vivo* (Darlington et al., 2000; Hao et al., 1997, 1999). In the late afternoon/early evening at ~ZT12, about 6–8 hours after their mRNAs, PER and TIM proteins begin to accumulate in the cytoplasm and form a protein complex that enter the nucleus in the middle of the night (Curtin et al., 1995; Lee et al., 1996; Price et al., 1998; Zeng et al., 1996). Accumulating level of PER-TIM protein complex binds CLK-CYC during the night to inhibit their transcriptional activity, and once PER and TIM are degraded early in the morning, the next round of CLK-CYC activation begins (Allada and Chung, 2010; Hardin, 2005, 2011; Fig I.2).

Another component of this loop is *cryptochrome* (*cry*), which encodes a blue light photoreceptor that also functions as a clock component in some tissues (Stanewsky et al.,

1998). CRY protein accumulates during the dark and declines during the day, driven by environmental light cycles. CRY binds directly to TIM in a light-dependent manner, which commits TIM to degradation in the proteasome mediated by the F-box protein JETLAG (JET) (Busza et al., 2004; Ceriani et al., 1999; Dissel et al., 2004; Koh et al., 2006, 2006; Naidoo et al., 1999; Peschel et al., 2009). CRY is also degraded in the proteasome upon activation by light, but CRY degradation occurs more slowly than TIM degradation (Busza et al., 2004; Dissel et al., 2004; Koh et al., 2006; Lin et al., 2001; Peschel et al., 2009).

I.C.2. Interlocked feedback loops

The interlocked transcriptional feedback loop is also regulated by the core feedback loop. Two CLK-CYC-dependent transcription factors, VRI and PDP1ε, and yet unidentified constitutive activator(s) mediate this second transcriptional feedback loop (Cyran et al., 2003; Glossop et al., 1999, 2003). In this loop, CLK-CYC activate transcription of *vri* and *Pdp1ε* between ~ ZT4 and ~ZT16 (Blau and Young, 1999; Cyran et al., 2003). *Clk* is constitutively activated by unknown activators independent of circadian oscillator function, as mutants that disrupt CLK–CYC transcriptional activity have constant high levels of *Clk* mRNA (Glossop et al., 1999).VRI protein accumulates in phase with *vri* mRNA, peaking in abundance at ~ ZT14. As VRI level increases, VRI binds D-box on the *Clk* promoter, thereby repressing *Clk* transcription (Cyran et al., 2003; Glossop et al., 2003). Another transcription factor, PDP1ε, rises to peak level at ~ ZT18, several hours after VRI level peak, also plays a role in *Clk* activation late at night,

but is not essential for *Clk* activation. (Benito et al., 2007; Cyran et al., 2003; Zheng et al., 2009). (Fig I.2).

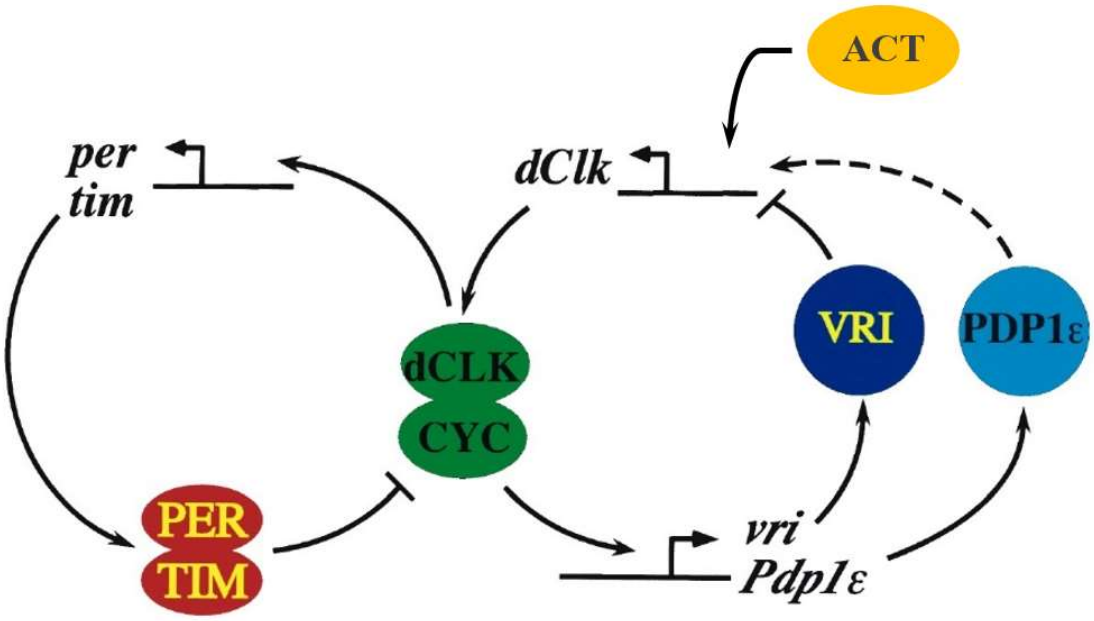


Fig I.2. Two transcription feedback loops of the *Drosophila* Clock (Adapted with permission from Cyran et al., 2003)

Two transcription feedback loops drive the *Drosophila* molecular clock. In the core loop, CLK-CYC directly activate transcription of *per* and *tim*. Inhibition of CLK-CYC activity is mediated by TIM/PER into the nucleus. In the second loop, CLK-CYC also activate *vri* and *Pdp1ε* transcription. *Clk* transcription is activated by unknown activator(s), and repressed by VRI, PDP1ε also plays a role on *Clk* transcription. Degradation of PER frees CLK-CYC to resume transcription of all the four target genes, thus restarting both loops simultaneously.

I.C.3. Posttranscriptional regulation of circadian transcription in *Drosophila*

Although the transcriptional feedback loops described earlier are the core of the molecular clock in *Drosophila*, they don't explain the ~24h oscillatory cycle as the whole process, which includes transcriptional activation, protein synthesis, nuclear localization, transcriptional repression, and repressor degradation, takes much less time to complete (Hardin, 2011). There is growing evidence suggesting that self-sustaining ~24 h oscillatory mechanisms are dependent on multiple regulatory pathways, such as temporal changes in the posttranslational regulation of core clock proteins. Posttranslational regulation of the inhibitory components imposes temporal delays between CLK-CYC transcriptional activation and PER-TIM repression. These temporal delays between activation and inhibition result in daily oscillations of CLK-CYC target gene transcription.

I.C.3.1 PER phosphorylation

PER is phosphorylated by DOUBLETIME (DBT), CASEIN KINASE 2 (CK2), SHAGGY (SGG) and NEMO (NMO). DBT binds to PER and promotes PER degradation via phosphorylation (Kim et al., 2007; Kloss et al., 1998, 2001; Price et al., 1998), whereas TIM binds to PER and prevents PER degradation (Kloss et al., 1998, 2001; Ko et al., 2002; Price et al., 1998). PER is progressively phosphorylated when complexed with DBT and CLK–CYC to repress transcription. A series of DBT and NMO mediated phosphorylation events result in a delay of DBT phosphorylation at PER serine 47, which is the final step that produces an atypical SLIMB binding site (Chiu et

al., 2008, 2011; Yu et al., 2011). The binding of SLIMB to PER triggers its degradation by the ubiquitin-proteasome pathway (Chiu et al., 2008), thus releasing repression of CLK-CYC, permitting a new cycle of transcriptional activation. PER is phosphorylated by CK2 at multiple kinase target sites to promote nuclear localization of PER-TIM complexes (Chiu et al., 2008; Lin et al., 2002a). Also, the phosphorylation of PER at multiple consensus proline-directed kinase target sites primes phosphorylation of S657 by SGG to promote PER nuclear localization instead of degradation (Chiu et al., 2008; Ko et al., 2010; Martinek et al., 2001). These temporal delays in PER nuclear transportation and degradation result in daily oscillations of CLK/CYC targets transcription in ~24h (Fig I.3).

I.C.3.2 CLK phosphorylation

Although *Clk* transcription is rhythmic, CLK protein levels remain constant (Houl et al., 2006; Yu et al., 2006). Thus, CLK-CYC-mediated transcription of target genes is dependent upon modifications such as phosphorylation and protein interactions rather than protein levels. CLK phosphorylation coincides with the entry of PER repression complexes into the nucleus followed by transcriptional repression (Kim and Edery, 2006; Menet et al., 2010; Yu et al., 2006, 2009). PER carries DBT into the nucleus (Kloss et al., 2001), while DBT is required for CLK phosphorylation (Kim and Edery, 2006; Yu et al., 2006, 2009). Interestingly, DBT doesn't phosphorylate CLK directly, however it must be present in the PER repression complex to mediate CLK phosphorylation by other kinases (Yu et al., 2009). The kinase responsible for CLK

phosphorylation remains to be identified, and one potential kinase implicated in CLK phosphorylation is NMO, as loss of *nmo* function increases CLK levels and shortens circadian period, and increasing *nmo* function decreases CLK levels and lengthens circadian period (Yu et al., 2011). CLK phosphorylation coincides with transcriptional repression (Kim and Edery, 2006; Yu et al., 2009, 2011), however it is not clear whether CLK phosphorylation is the cause for repression.

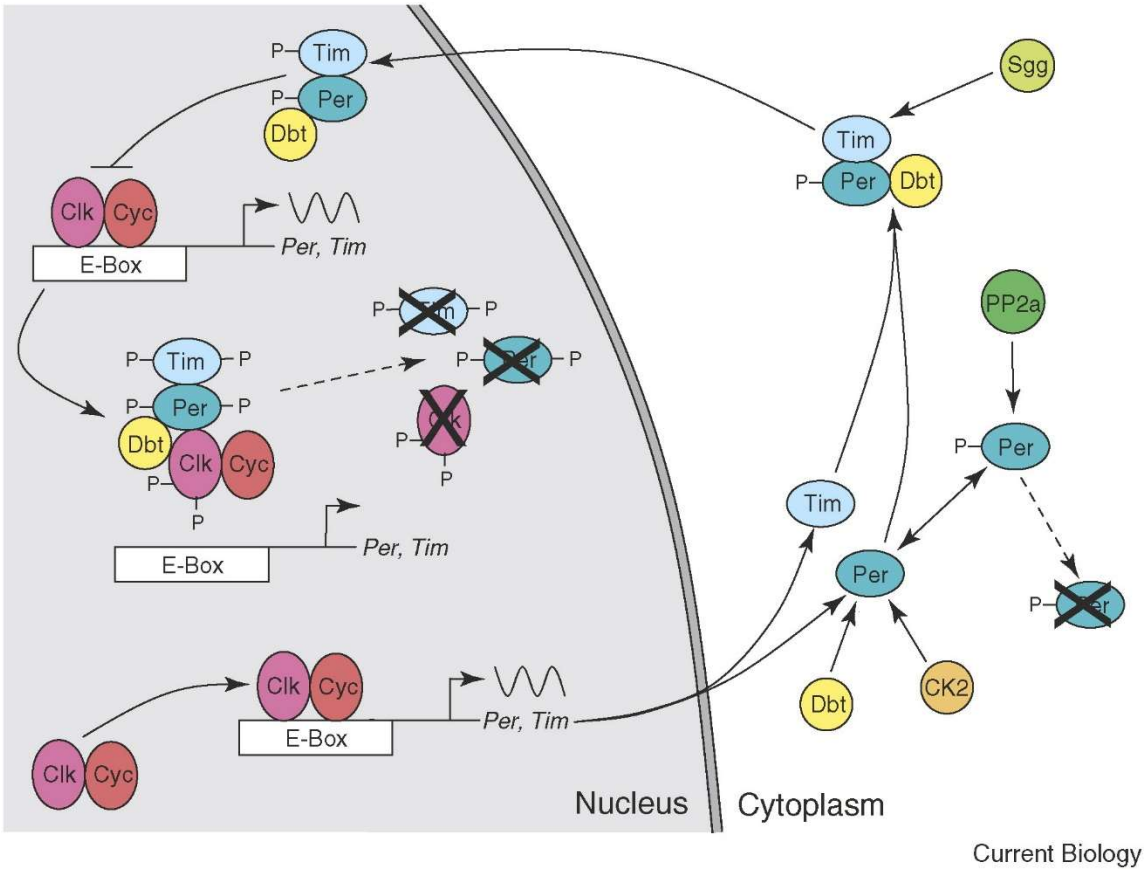


Fig I.3. Posttranscriptional regulation of the Per/Tim feedback loop (Taken with permission from Hardin, 2005)

PER is phosphorylated by DBT and CK2, which leads to its degradation. TIM binds to phosphorylated PER and stabilizes the protein. PER is also stabilized by PP2a through

dephosphorylation. SGG further phosphorylates PER/TIM and promotes their nucleus entry and transcriptional repression. PER and CLK are then destabilized via DBT or DBT bridged phosphorylation, and degraded. This allows the accumulation of non-phosphorylated (or hypo-phosphorylated) CLK to start another cycle of transcription.

I.D. Transcription repression in circadian feedback loops in *Drosophila*

A negative feedback loop occurs in biology when the product of a reaction leads to a decrease in that reaction. In this way, for any negative feedback loop, a repression step is critical for the stability and the precise timing of the system. Therefore, understanding the mechanism of transcriptional repression is essential for deciphering the regulation of the circadian feedback loops in *Drosophila*.

The timing of transcriptional events in the feedback loop coincide with, and are controlled by, rhythms in CLK-CYC binding to E-boxes (Yu et al., 2006). PER was previously found to inhibit CLK-CYC binding to E-boxes *in vitro* (Lee et al., 1999), which suggests that the rhythmic transcription of CLK target genes are mediated by PER-dependent rhythms in E-box binding by CLK-CYC. Chromatin immunoprecipitation (ChIP) experiments using fly heads show that CLK-CYC rhythmically bind E-boxes in the *per* CRS and the *tim* upstream sequence (Yu et al., 2006; Fig I.4 A, B). PER is required for the rhythmic binding of CLK complexes, as CLK constantly binds to *per* and *tim* promoters in per^{01} mutant flies (Yu et al., 2006; Fig I.4 C, D), indicating that PER inhibits transcription by removing CLK-CYC from E-boxes during repression phase *in vivo*. However, the mechanism by which CLK-CYC heterodimers are removed from E-boxes during repression is not well understood.

Interestingly, co-expression of another transcription factor CLOCKWORK ORANGE (CWO), strongly enhances PER-mediated repression in cultured *Drosophila* Schneider 2 (S2) cells (Kadener et al., 2007), suggesting that PER by itself is unable to efficiently remove CLK from DNA in the absence of other transcription repressors. Thus, identifying and characterizing the molecular function of additional clock genes, especially components that involve in the repression process, are important to improve our understanding of the circadian transcriptional regulatory mechanism.

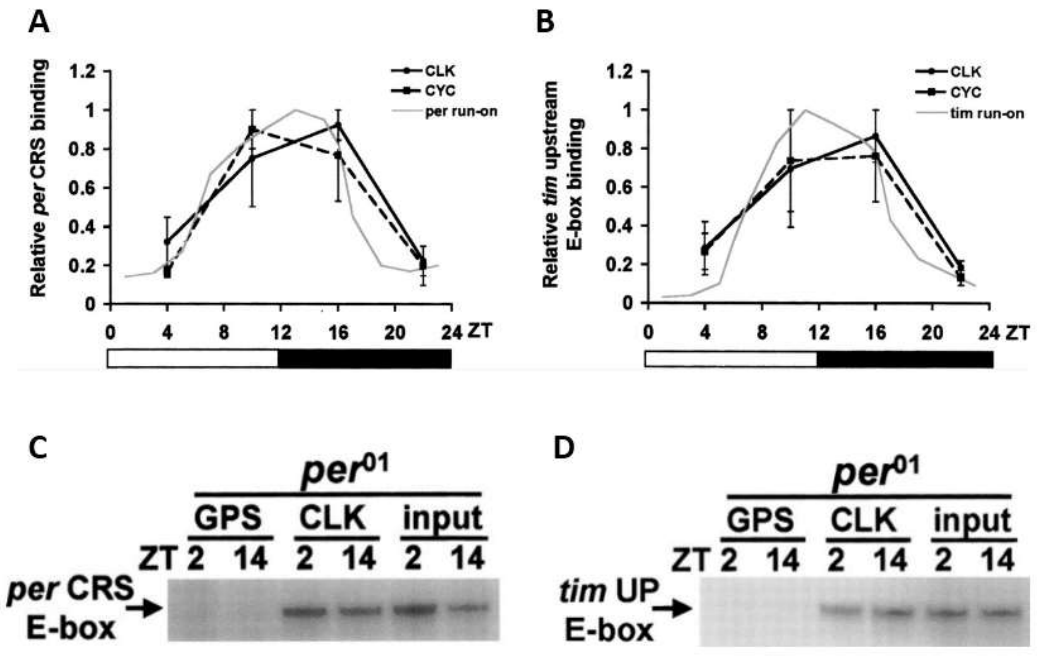


Fig I.4. PER dependent rhythmic binding of CLK on E-boxes (Adapted with permission from Yu et al., 2006)

(A, B) ChIP assays indicated CLK-CYC heterodimers rhythmically bind E-boxes on **(A)** *per* and **(B)** *tim* promoters *in vivo*. **(C, D)** ChIP assays were performed on per^{01} flies. In the absence of PER, the cycling of CLK binding to **(C)** *per* and **(D)** *tim* promoters is abolished *in vivo*.

I.E. CWO is a circadian transcriptional regulator in *Drosophila*

In 2007, three groups had independently identified CWO as a new clock component (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007). Since the main focus of this study is to investigate the function of CWO in the circadian clock of *Drosophila*, here I will briefly summarize the background information supporting CWO function as a circadian transcriptional regulator that was previously described in the earlier publications.

I.E.1. *cwo* transcription is controlled by CLK through canonical E-boxes

Several studies had identified *cwo* as a rhythmically expressed gene, whose mRNA was reported to oscillate in a circadian manner by previous microarray and RNase protection assays (Claridge-Chang et al., 2001; Lin et al., 2002b; McDonald et al., 2001; Ueda et al., 2002). The temporal expression of *cwo* mRNA rhythmically changes in the light-dark condition (LD) and constant darkness (DD), peaking closely in phase with other CLK target genes *per* and *tim*, though having a noticeably lower amplitude (Fig I.5A). This cycling of *cwo* mRNA is abolished in Clk^{Jrk} mutants, which lacks its activation domain on CLK, suggesting that *cwo* is regulated by the same molecular mechanism as *per* and *tim* (Kadener et al., 2007; Lim et al., 2007; Fig I.5B). There are six E-boxes within the promoter of *cwo* (2 kb upstream of the transcriptional start site), and 15 E-boxes within the first intron (Kadener et al., 2007), suggesting that it is a CLK target, as E-boxes are necessary for transcriptional oscillations and have been shown to mediate CLK activation followed by PER repression (McDonald and Rosbash, 2001; Wang et

al., 2001). This was later confirmed by genome-wide ChIP-on-chip and ChIP-seq analysis that aiming to identify direct CLK-CYC targets, and *cwo* was among the top on the list of all other CLK targets (Abruzzi et al., 2011; Menet et al., 2010).

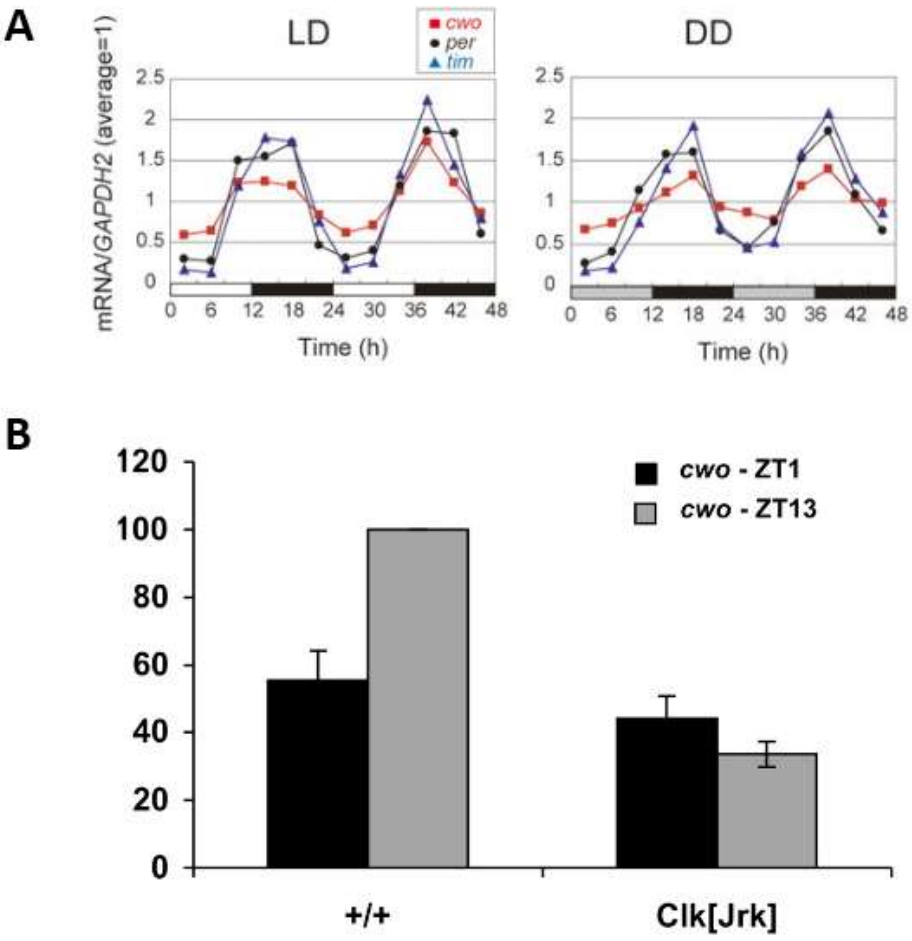


Fig I.5. *cwo* transcription is rhythmically activated by CLK (Adapted with permission from Lim et al., 2007; Matsumoto et al., 2007).

(A) Temporal expression profiles of *cwo* (red square), *per* (black circle), and *tim* (blue triangle) mRNA in wild-type flies under LD and DD. *cwo* mRNA is in phase with *per* and *tim* mRNA. **(B)** RT-PCR analysis of *cwo* transcript levels in wild-type and Clk^{Jrk} mutants at ZT1 and ZT13. *cwo* mRNA rhythm is abolished in Clk^{Jrk} mutant.

I.E.2. *cwo* encodes a transcriptional factor that specifically binds E-boxes *in vitro*

cwo encodes a 685 amino acid protein, contains bHLH and ORANGE domains, which indicates that CWO is a candidate transcription factor (Davis and Turner, 2001). The CWO protein, however, lacks the tetrapeptide domain WRPW, a 4-amino-acid transcription repression domain that generally exists at the C-terminus of Hey, hairy, or E(spl) subfamilies in the bHLH-ORANGE family (Davis and Turner, 2001; Fisher et al., 1996), and belongs to the Stra13 subfamily. Previous ChIP-on-chip assays reveal that CWO protein binds to canonical E-boxes in S2 cells. Among the 1512 binding sites detected in S2 cells, CWO protein binds to the known clock genes *vri*, *Pdp1* and *cwo* itself, all of which have the E-box sequences in their promoter region (Matsumoto et al., 2007; Fig I.6). Gel-shift analyses indicates specific binding of CWO to a CACGTG E-box but not mutant E-box probes, and this binding is partially competed by an unlabeled E-box, but not a mutated E-box fragment, and is super shifted by GST antibodies (Lim et al., 2007). All these results, taken together, indicate that CWO is a transcription factor and specifically binds E-boxes *in vitro*.

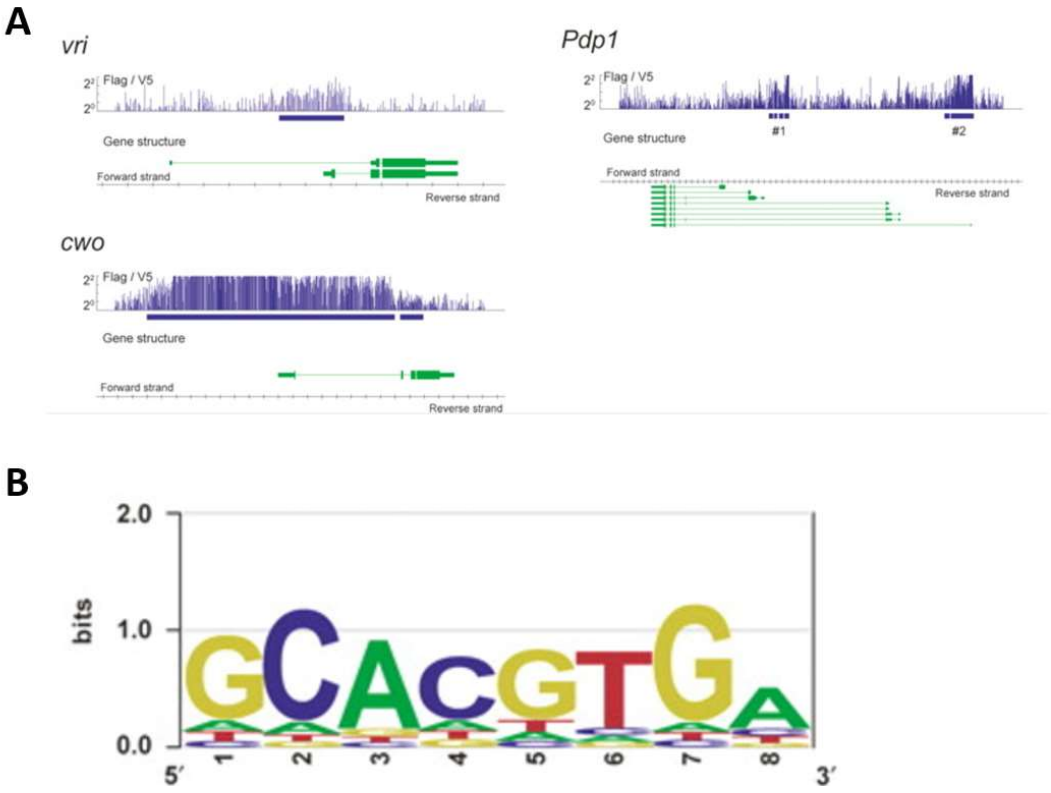


Fig I.6. CWO protein directly targets known clock genes (Taken with permission from Matsumoto et al., 2007).

(A) ChIP-on-chip assays reveal CWO binding to the known clock genes *vri*, *Pdp1* and *cwo*. **(B)** Bioinformatic search for the consensus DNA sequence identified a sequence recognized by the CWO protein containing the canonical circadian E-box (CACGTG).

I.E.3. *cwo* is a core clock component and negatively regulates the expression of clock genes

Various *cwo*-deficient strains, including two strains that contain unique transposon insertions at the beginning and at the end of the first intron (cwo^{5073} and cwo^{4027}), one ethane methyl sulfonate (EMS) mutagenized strain (cwo^{B9}), and several *cwo* RNAi knockdown strains, have strong circadian locomotor activity phenotypes (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). During the constant darkness, more than half of these *cwo* RNAi knockdown or mutant flies are arrhythmic, and the remaining flies have weaker and longer period rhythms than control strains. Along with these behavior data, CWO proteins were specifically detected in the oscillator neurons by immunostaining and enhancer trap experiments (Kadener et al., 2007; Matsumoto et al., 2007; Shafer et al., 2006). Importantly, in homozygous cwo^{5073} mutants, CWO is not detected in the brain oscillator neurons (Matsumoto et al., 2007). These data suggest that *cwo* is a critical clock component and functions in brain oscillator neurons to maintain behavioral rhythms.

At the molecular level, CWO was reported to negatively regulate the expression of clock genes. In cultured S2 cells, overexpression of CWO reduces the basal transcription of *per*, *tim*, *vri* and *Pdp1ε* promoter-driven luciferase reporter genes (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007). In addition, with PER present, CWO represses CLK mediated transcription 5 to 10 fold, and the expression of a fixed amount of CWO strongly enhances PER-mediated repression, indicating that CWO is a strong transcription repressor that can cooperate with PER to repress CLK-

CYC-mediated transcription in S2 cells (Kadener et al., 2007; Matsumoto et al., 2007; Fig I.7A, B). This suppression is also observed *in vivo*: the troughs of *per*, *tim*, *vri*, and *Pdp1* mRNA during early morning at ZT3 are significantly elevated by 2 to 3 fold in *cwo* RNAi knockdown or mutant strains relative to those of wild-type flies (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008; Fig I.7C). Interestingly, in these *cwo*-deficient strains, reduced transcript levels of the same genes were obtained during the peak at ~ ZT13 (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008), indicating CWO also potentially acts as a transcriptional activator to promote the transcription of these target genes. The increased trough and decreased peak mRNA levels thus reduce the amplitude of expression of these genes to half the level found in wild-type flies under LD conditions. This reduced amplitude of circadian oscillators in *cwo* RNAi transgenic and *cwo* mutant flies was also confirmed under DD conditions (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007), suggesting that *cwo* functions to produce a high-amplitude oscillations in clock genes expression.

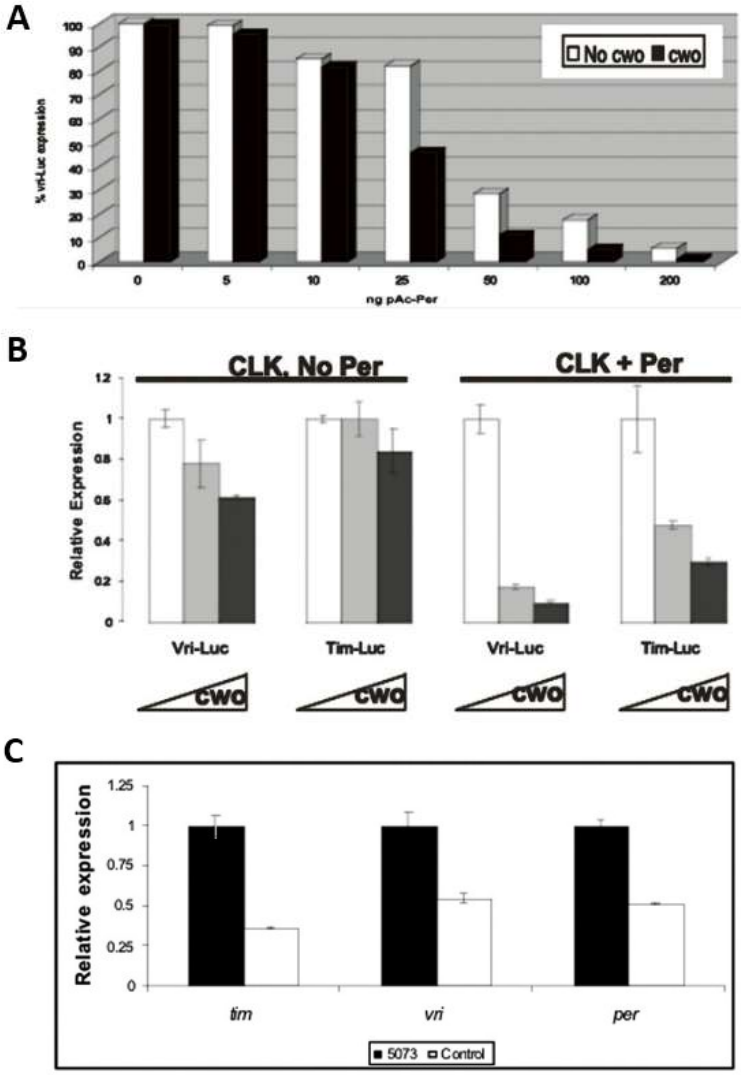


Fig I.7. CWO represses the expression of clock genes in cell culture and *in vivo* (Taken with permission from Kadener et al., 2007).

(A) PER repression of CLK-mediated transcription in the presence or absence of CWO in S2 cell culture. **(B)** CWO repression of CLK-mediated transcription in the presence or absence of PER in S2 cell culture. **(C)** Comparison of trough values for *tim*, *vri*, and *per* for wild-type and *cwo*^{5073} flies measured by oligonucleotide microarray.

I.F. The goal of the study

The rhythmic binding of CLK-CYC to E-boxes is a key step in regulating the core and interlocked feedback loops. As described earlier, PER is required for removal of CLK-CYC from the DNA during repression phase (Yu et al., 2006). However, the detailed mechanism of this PER-mediated transcriptional repression process still remains unknown. Previous experiments in cell culture indicated that CWO binds to the E-box on the promoter region of clock genes and represses their expression, suggesting that CWO is another transcription repressor in the circadian feedback loops in *Drosophila*. However, *in vivo* CWO binding data is lacking to further support this hypothesis. Does CWO bind the promoter of CLK target genes *in vivo*, like that observed in cell culture experiments? If CWO binds these clock genes *in vivo*, does its binding change in a temporal manner similar to the case of CLK binding, or does CWO constantly bind to those genes? If CWO binds to CLK binding sties on the promoter regions, does CWO binding affect CLK binding, thus playing a role in the circadian transcriptional regulation? By carrying out *in vivo* ChIP assays and carefully compare the binding of CWO to the binding pattern of CLK, I will be able to address these important questions, which will allow me to further investigate the function of CWO during transcriptional repression, and potentially provide a deeper understanding of the detailed mechanism of transactional repression in the circadian feedback loops. Importantly, CWO has two mammalian orthologs DEC1 and DEC2, which also target E-boxes bound by CLOCK-BMAL1 and repress gene expression *in vitro* (Honma et al., 2002a; Kawamoto et al., 2004; LI et al., 2004a), suggesting conserved transcriptional repression mechanism

between *Drosophila* and mammals. Therefore, studying the function of CWO and its transcriptional regulatory mechanism may shed light onto our understanding of the potentially similar mechanism that operate in the mammalian clock.

Previous ChIP-on-chip assays showed that CWO binds 1512 sites in S2 cell culture (Matsumoto et al., 2007). However, in that experiment, CWO protein was overexpressed, leading to potential false-positive interactions that CWO may bind to many other targets that it would not bind endogenously, since the concentration of the induced CWO protein could be hundreds of fold higher than endogenous protein. Moreover, the binding sites of CWO detected in cell culture does not necessarily fully reflect CWO binding in *Drosophila.* Indeed, the known clock genes *vri* and *Pdp1*, which have the E-box sequences in their promoter region were detected as CWO targets; however, other clock genes, *per* and *tim* that also have the E-box sequences were not detected in the experiment. These results raise the possibility that there could be false-positives and false-negatives in this cell culture assay which make it less informative. Consequently, carrying out ChIP-seq experiment *in vivo* would be critical to identify potential CWO targets and to characterize the pattern of CWO binding at genome-wide level. Combining ChIP-seq with RNA-seq analysis, I could further characterize how CWO regulates the expression of clock genes and other potential output genes globally.

CHAPTER II

MATERIALS AND METHODS

II.A. Transgene construction and transgenic fly generation

DNA fragments containing wild-type or mutant E-boxes from the upstream *tim* circadian enhancer were used to construct GFP-reporter transgenes. These 136bp fragments extend from -578 to -714 relative to the *tim* transcription start site, and contain "E1-E2" E-box motifs that are wild-type (E1-E2), E1 mutant (mE1-E2), E2 mutant (E1-mE2) or E1-E2 mutant (mE1-mE2). These wild-type and mutant E-box fragments were generated by PCR amplification using the following primer sets: E1-E2, 5'-CACCTTTGGCAAATAAACGTGCGGCA-3' and 5'-TGCCGGCGTTTGTGCGAA-3'; mE1-E2, 5'-CACCTTTGGCAAATAAACGTGCGGCACGTTGTGATTAAGATCTAGCCGAT-3' and 5'-TGCCGGCGTTTGTGCGAA-3'; E1-mE2, 5'-CACCTTTGGCAAATAAGATCTCGGAGATTTGTGATTACACGTGAGCCGAT-3' and 5'-TGCCGGCGTTTGTGCGAA-3'; mE1-mE2, 5'-CACCTTTGGCAAATAAGATCTCGGAGATTTGTGATTAAGATCTAGCCGAT-3' and 5'-TGCCGGCGTTTGTGCGAA-3'. The PCR products were inserted into the pENTR/D-TOPO vector using pENTR Directional TOPO cloning kit (Invitrogen), and then subcloned into the pHPdesteGFP vector, which expresses Green Fluorescent Protein (GFP) according to the enhancer sequence inserted [27], using Gateway LR-Clonase System (Invitrogen). The nucleotide sequences of all transgenes were confirmed

by sequencing. The resulting transgenes were injected into embryos (BestGene) for recombination into the attp18 genomic site via PhiC31-mediated transgenesis to yield *tim* circadian enhancer GFP (*tim*-CEG) flies (Bischof and Basler, 2008; Groth et al., 2004; Venken et al., 2006, 2008).

A C-terminal c-Myc and 3xHA tagged *cwo* transgene (*cwo*-HA) was constructed via recombineering. High Fidelity DNA polymerase (Invitrogen) was used to amplify the Frt-ampicillin-Frt (Frt-Amp-Frt) cassette from FRT-gb2-amp-FRT plasmid (Gene Bridges) using primer *cwo*-MyC-3xHA-L 5'gcagcggtggctaaggccaaactggagcaggccatgaaccagagctggGAACAAAAACTTATTTCTG AAGAAGATCTGaatagcgccgtcgacTACCCATACGACGTACCAGATTACGCTTACCC ATACGACGTACCAGATTACGCTTACCCATACGACGTACCAGATTACGCTtagGCA GCCCAATTCCGATCATATTC-3' (53 nucleotides (nts) of *cwo* sequence upstream of the stop codon (lowercase), 30 nts of the c-Myc sequence (uppercase), 15 nts of the linker sequence (lowercase), 81 nts of the 3xHA sequence (lowpercase), a stop codon (uppercase) and 23 nts of the Frt-Amp-Frt cassette (lowercase)), and cwo-R 5′tactgaggtagtgttgttccatctgtcgacccattgcattgcgattgctttgcTGGATCCCCTCGAGGGACCT AT-3′ (53 nts of *cwo* sequence downstream of the stop codon (lowercase), and 23 nts from the 3′ end of Frt-Amp-Frt cassette (lowercase)). This PCR reaction was run at melting temperature (TM) 56°C for 35 cycles, treated with DpnI enzyme and purified. This fragment was used to transform SW102 cells harboring the BAC clone CH321-18B09 (BAC-PAC Resources Center), which contains the *cwo* genomic region 12.494 kb, and recombinants containing the Frt-Amp-Frt cassette inserted into *cwo* were

selected on plates containing ampicillin. The ampicillin gene was removed by inducing recombination at the Frt sites (Venken et al., 2008), resulting in the chloramphenicol-resistant *cwo*-Myc-HA p(ACMAN) clone. *cwo*-Myc-HA was sequenced to confirm the C-terminal Myc-HA tag fusion. The *cwo*-Myc-HA transgene was then inserted into attP40 on chromosome 2 via PhiC31-mediated transgenesis (Groth et al., 2004).

II.B. ChIP Protocol *

Step 1. Isolating fly heads

In this step, the procedure for isolating fly heads is described. Fly heads are typically used to study molecular clock mechanisms in *Drosophila* because they are highly enriched for clock cells. The vast majority of clock cells in fly heads are photoreceptors, which show rhythms in clock protein expression similar to brain pacemaker neurons and peripheral tissues (Glossop & Hardin, 2002). Although it would be useful to assess DNA binding profiles of feedback loop components in individual tissues of flies, it is not practical to purify large quantities of fly tissues in contrast to the situation in mammals.

Fly Collection:

1. Place 50ml Falcon tubes and a 80mm funnel on dry ice for cooling.
2. Collect approximately 20ml of flies into a Falcon tube on dry ice using the funnel.
3. Freeze flies at -80°C for at least 3 hours before collecting heads.

*** 1. This chapter is reprinted with permission from Zhou J, Yu W, Hardin PE (2015) ChIPping Away at the Drosophila Clock. Methods in Enzymology. Volume 551, Pages 323-347. Copyright [2015] by Elsevier. The solutions needed for ChIP are listed in the Appendix at the end of this dissertation.**

4. You may keep flies frozen at -80°C for many weeks or months if necessary before collecting fly heads.

Head Collection:

1. Place a #25 and a #40 sieve both with the collection holders and a 60mm funnel on dry ice for cooling.
2. Vortex each Falcon tube for 20 seconds twice, then shake the tube vigorously for 20 seconds, keep the tube on dry ice for approximately 1-2 minutes in between vortexing and shaking.
3. Stack the #40 sieve on top of the collection pan and pour the entire 20ml of flies from the Falcon tube onto the #40 sieve. Using a soft paint brush, brush the flies so that the wings and legs fall through the #40 sieve. Fly heads and bodies will be left on the #40 sieve.
4. Stack the #25 sieve on top of another collection pan and pour the fly heads and bodies from the #40 sieve onto the #25 sieve, brush the heads and bodies on the #25 sieve with a hard paint brush until all the heads fall through and into the collection pan.
5. Transfer fly heads from the collection pan into an Eppendorf (EP) tube that was pre-cooled on dry ice using the 60mm funnel. Label each EP tube and keep at -80°C.
6. The procedure can be stopped here when the samples are frozen at -80°C, or continue to the next step.

Step 2. X-Nuclei preparation

This step describes how to prepare cross-linked chromatin-protein complexes. Because fly heads are covered with cuticle, it is important to first grind the heads using a homogenizer to break apart the cuticle so that nuclei can be efficiently released by homogenization buffer. During the homogenization process, any protein-DNA complexes present in cell nuclei will be stabilized by formaldehyde cross-linking.

Preparing cross-linked chromatin-protein complexes:

1. Measure 1ml fly heads.
2. Warm up 2x XIP-HB-HSEEIT, 10x XIP-TSEE, and 1.4M glycine solutions at 37°C water bath.
3. Prepare 5ml of Homogenization buffer (HB) at room temperature (RT), 7ml of Homogenate Dilution buffer (HDB) and 8ml wash buffer for each sample.
4. Place a 7ml Homogenizer with loose pestle and a 60mm funnel on dry ice.
5. Place 1ml frozen fly heads into a 7ml homogenizer and grind them on dry ice for 80 strokes.
6. Pour ground heads into another 7ml Homogenizer at RT or a 25°C water bath using the dry ice-cooled funnel.
7. Immediately add 5ml of HB, gently homogenize for 10 minutes at RT with occasional vortexing.
8. Add 570µl of 1.4M Glycine, homogenize gently for 5 minutes at RT with occasional vortexing.

9. Filter the homogenate through 100µm nylon mesh into a 50ml Falcon tube by placing the mesh on top of the Falcon tube and pouring the homogenate onto the mesh.
10. Wash the Homogenizer using 7ml of HDB by homogenizing for 3-5 strokes, then filter the homogenate through the nylon mesh.
11. Pellet X-Nuclei
    a) Spin the 50ml Falcon tube at 1,500 x *g* for 5 minutes, carefully remove the supernatant, do not touch the pellet at the bottom of the tube.
    b) Re-suspend the pellet with 1ml wash buffer, and transfer the suspension into 3 EP tubes evenly.
    c) Centrifuge the EP tubes at 950 x *g* for 5 minutes, remove supernatant.
    d) Wash the pellets twice with 1ml wash buffer, each time spin the tube at 950 x *g* for 5 minutes.
    e) Transfer all nuclei suspension into one EP tube
    f) Spin at 950 x *g* for 5 minutes, remove the supernatant
    g) Label the tube containing X-Nuclei, Store at -80°C.
    h) The procedure can be stopped at this point, if necessary, once the sample is frozen at -80°C.

Step 3. Sonication

In this step, the DNA in cross-linked protein-DNA complexes will be sheared into short fragments. Chromatin shearing is a critical step because DNA fragment size will

significantly affect IP efficiency and background. DNA shearing protocols differ depending on the type of sonicator that is used. Here I provide two protocols for DNA shearing using a standard Microson sonicator or a Diagenode Bioruptor sonicator. Since the volume to be sonicated is low (~400µl), a 3/32 inch microprobe should be used in the Microson sonicator. After the DNA is sonicated, the size of sheared DNA is measured (Fig. 2), and should ideally fall into the 200bp-800bp size range.

Microson sonication:

1. Thaw X-Nuclei on ice. Add 3x volume of XIP-SonicBuffer to X-Nuclei. Typically the volume of nuclei is ~100µl, thus the sonication volume will be ~400µl.
2. Set the Microson XL 2000 sonicator output at 4-5 watts on the display.
3. Sonicate for 10 seconds x 15 times in a cold ethanol bath on crushed ice, 150 seconds in total. Wait 50 seconds between each sonication.
4. Centrifuge at 25,000 x *g* for 10 minutes at 4°C to remove debris (most of the debris is cuticle), and save supernatant as X-Nuclear extract (SXN).
5. Estimate the concentration of SXN
    a) Make a series of standards by diluting bovine serum albumin (BSA) into XIP-SonicBuffer.
    b) Dilute 2ml of Bio-Rad Protein Assay Dye Reagent concentrate (Catalog #500-0006) with 8ml of H_2O to make protein dye mix.
    c) Mix 5µl of each standard or sample with 1ml of protein dye mix by vortexing.
    d) Measure each standard in a spectrophotometer and generate a standard curve.

f) Measure each SXN sample in a spectrophotometer, calculate the protein concentration of each sample according to the standard curve.

6. Calculate the volume of SXN needed for ChIP (500µg protein) and input (50µg protein), and store aliquots at -80°C.
7. Stop here if necessary with the sample frozen at -80°C, or continue to the next step.

Diagenode Bioruptor sonication:

1. Resuspend X-Nuclei (~100µl) in 400µl of Bioruptor sonication buffer.
2. Split the sample into two Bioruptor tubes (250µl in each).
3. Sonicate using the following program:

   A cycle of 30 seconds on and 30 seconds off is repeated for a total of 20-30 cycles. Vortex the tubes every 10 cycles.
4. Centrifuge at 15,000 x *g* for 10 minutes at 4°C to remove debris, and save supernatant as Bioruptor X-Nuclear extract (BXN).
5. Estimate concentration and calculate the volume for ChIP using the same method as described in steps 5 and 6 for the Microson sonicator. Store aliquots at -80°C.
6. Stop here if necessary with the sample frozen at -80°C, or continue to the next step.

Determine DNA fragment size:

1. Add 2µl of 5M NaCl to a 10-20µl aliquot of the SXN or BXN and incubate in a 65°C water bath for 6 hours to overnight (or incubate in a boiling water bath for 15 minutes) to reverse the crosslinks, and centrifuge at max speed for 5 minutes.

2. Run the supernatant on a 2% agarose gel to determine the fragment size**.**

Step 4. IP and washes

In this step, the Protein-DNA complexes will be incubated with an antibody raised against the protein of interest to form antibody-protein-DNA complex. The antibody-protein-DNA complex will be isolated using Dynabeads coupled to either Protein A or Protein G, which are bacterial proteins with high affinity for immunoglobulins. Once the antibody-protein-DNA complexes are bound to Dynabeads, non-specific protein-DNA complexes and free DNA fragments are removed during a series of washes, leaving specific antibody-protein-DNA complexes bound to the Dynabeads.

Isolating antibody-protein-DNA complexes:

1. Pre-incubate SXN or BXN with antibody.
    a) Add 3µl antibody to the SXN or BXN (500µg), then add IP buffer to SXN samples to bring the total volume to ~800µl or add IP buffer to BXN samples to dilute the sample 10-fold (~2.5ml).
    b) Add 10% NaN_3 to a final concentration of 0.025%.
    c) Place the sample in a tube rotator set at ~10 rpm overnight at 4°C.
2. Blocking beads (prepare the same day).
    a) Use 30-50ul Dynabeads for each sample. Wash Dynabeads with 1ml IP buffer. For this and subsequent washes, capture the beads using a magnetic stand, add the solution to resuspend the beads, and then recapture beads.
    b) Wash the beads with 1ml blocking solution twice.

c) Re-suspend the beads with blocking buffer, and add NaN_3 to a final concentration of 0.025%.

d) Rotate the beads at ~10 rpm overnight at 4° C.

3. Incubate immunocomplexes with Dynabeads.

a) Capture blocked beads, wash beads with 1ml IP buffer.

b) Re-suspend the beads with the pre-incubated SXN or BXN with antibody.

c) Rotate at ~10 rpm at 4°C for 2 hours.

4. Wash (Rotate at ~10 rpm for 5 minutes at 4°C for each wash).

a) Wash with 1ml of IP buffer twice.

b) Wash with 1ml of LowSalt buffer twice.

c) Wash with 1ml of HiSalt buffer twice.

d) Wash with 1ml of Li buffer twice.

e) Wash with 1ml of TE buffer twice.

Step 5. Elution and DNA extraction

In this step, the Antibody-Protein-DNA complex is eluted from the beads, treated to remove RNA and proteins, and then reverse-cross-linked. DNA fragments are then purified for Real-Time quantitative PCR (hereafter qPCR) or sequencing analysis.

Isolating DNA from immunoprecipitates:

1. Elution

a) Add 50µl of elution buffer to the washed beads and incubate at 65°C for 15 minutes (vortex every 2-3 minutes), then move the supernatant (50µl) to a new tube.

b) Repeat the elution on the same beads using another 50µl elution buffer, combine the supernatants together to give a total volume of 100µl.

c) Stop here if necessary with the sample frozen at -80°C, or continue to the next step.

2. DNA extraction:

a) Add one volume (100µl) of 2xTE buffer to the eluates. For input samples, add 1x TE buffer to total volume of 200µl.

b) Add RNase A to a final concentration of 50µg/ml, and incubate at 37°C for 30 minutes.

c) Adjust SDS in the input sample to 0.5%, then add Proteinase K to all the samples to a final concentration of 1µg /µl. Incubate at 37°C for 6 hours or overnight.

d) Add 5M NaCl to final concentration of 0.3M, reverse cross-link at 65°C for 6 hours to overnight.

e) Add 300µl of phenol chloroform for extraction, vortex, centrifuge for 5 minutes at 14000 rpm at 4°C. Remove and save the upper layer into another labeled EP tube. Add 50µl of 1xTE to the lower layer, vortex, and centrifuge for 5 minutes at 14000 rpm at 4°C. Remove the upper layer (~50µl) and combine with the previously extracted upper layer.

f) To prepare DNA for qPCR analysis, add 650ml of ethanol, glycogen (0.5µl of 20µg/µl) and 0.5µg sonicated salmon sperm DNA to the IP sample (do not add salmon sperm DNA to the input sample, only glycogen), and precipitate at -20°C overnight. To prepare DNA for ChIP-seq, do not add salmon sperm DNA, add 1µg glycogen in total instead.

g) Centrifuge at 14000 rpm for 30 minutes at 4°C, remove the supernatant. Add 1ml of cold 70% ethanol, centrifuge at 14000 rpm for 10 minutes at 4°C, air-dry for 10 min, then re-suspend in 50µl of 1xTE for qPCR or ChIP-seq.

h) Stop here if necessary with the sample frozen at -80°C, or continue to the next step.

Step 6. qPCR analysis

In this step, qPCR is carried out to quantify the amount of DNA that was immunoprecipitated at a particular target site in each sample, which is a measurement of the affinity of the protein for that target site. This method quantifies binding at a known or hypothesized site as a percentage of this site in input DNA (% of input) minus the % of input value for a negative control site that shows no binding. Alternatively, samples can be used to prepare libraries for ChIP-seq analysis, which will be discussed in more general terms below.

Quantifying immunoprecipitated DNA:

1. Design two sets of primers: one set for amplifying a 100-200 bp DNA fragment containing a target binding site, the other set for amplifying an untranslated region

region or other genomic region that is known not to be bound by the protein of interest as a background control.

2. There are several methods to quantify DNA levels using qPCR. Here I will use the standard curve method. To make a standard curve for the qPCR, a series of dilutions for one of the inputs needs to be generated using TE buffer. Set the standard quantity as follows:

Dilution of the input sample	Arbitrary quantity of standard
1:100,000	1
1:20,000	5
1:10,000	10
1:2,000	50
1:1,000	100
1:200	500
1:100	1000

3. Dilute each input sample into 1:1000 and each IP sample into 1:10 with TE buffer to use as DNA templates.

4. Reaction setup in 96-well PCR plates (Cat.# MLL9601, BIO-RAD):

7.5 µl	Ssofast EvaGreen Supermix (Cat. #172-5201, BIO-RAD)
1.25 µl	10uM Primer Forward
1.25 µl	10uM Primer Reverse

5 µl	Diluted DNA template

5. Centrifuge the 96 well plates in 1500 rpm for 2 minutes, then put the plate in a BIO-RAD CFX96 real-time PCR machine, design the qPCR program according to the fragment length and primer annealing temperature. Save the data file containing the qPCR results, and analyze the data with CFX manager software according to the manufacturer's instructions.
6. A standard curve will be automatically generated from the dilutions of the input sample (see #2 in this section) by the CFX manager software. The qPCR Starting Quantity (SQ) for each IP sample and input will then be calculated by the software based on the standard curve that was produced. Because the IP samples and the inputs were diluted before qPCR, the original quantity of IP samples and inputs is then calculated:

   IP Quantity = IP SQ x 10, and input Quantity = input SQ x 1000 (the 10 and 1000 multipliers come from the dilution fold in step 3). If different IP and input dilutions were used, the equation should be adjusted accordingly.
7. Relative ChIP abundance is represented as the % of input, which is the proportion of DNA fragments that are enriched from the starting material (input). The % of input is calculated as follows:

   % of input = (IP Quantity) / (input Quantity x 10) x 100% (the 10 multiplier comes from the 500µg used for IP versus the 50µg for input. The same calculation is used for negative control data.

8. Correction for non-specific binding = % of input for binding site - % of input for negative control.

Primers used for ChIP-qPCR

CLK and CWO binding to E-boxes upstream of *tim*, *per*, *vri,* the rhythmically expressed *Pdp1* promoter from wild-type flies and the upstream *tim* circadian enhancer in *tim*-CEG flies were quantified via qPCR, and corrected for nonspecific binding to an intergenic region on chromosome 3R (nucleotides 29576172 to 29576303).

The primers used for ChIP-qPCR were as follows: for *tim* E-boxes, 5'-ACACTGACCGAAACACCCACTC-3' and 5'-GCGGCACGTTGTGATTACACG-3'; for *per* E-boxes, 5'-GGGTGAGTAATGCCGTTGCGAAAT-3' and 5'-ATTTGCTGGCCAAGTCACGCAGTT-3'; for *vri* E-boxes, 5'-CTGGTGCCTCACATTCCACG-3' and 5'- CAGCAGTCAAGTTATAGCAGCGC-3'; for *Pdp1* E-boxes, 5'-GCACTCTCATTCTCTCTGTCGC-3' and 5'-ACTTGGGGGACTGGAACTG-3'; for *tim*-CEG, 5'-GCCCCCTTCACCTTTGGCAAATA-3' and 5'-TACAAGAAAGCTGGGTCGGCG-3'; and for the intergenic region, 5'-CAGGAGTCGVAGGACCAACC-3' and 5'-GTCCTGAGAGGCTGAGAGGC-3'.

II.C. Quantitative RT-PCR

Quantitative RT-PCR was performed as described in II.B.6, with some modifications, to extract total RNA and measure *GFP* mRNA levels. Total RNA was isolated from frozen

fly heads using Trizol (Invitrogen), and treated with a Turbo DNase DNA-free kit (Ambion) to eliminate genomic DNA contamination. DNA-free total RNA (1.0 μg) was reverse transcribed using oligo (dT) 12–28 primers (Invitrogen) and Superscript II (Invitrogen). The reverse transcription (RT) product was amplified with SsoFast qPCR Supermix (Bio-Rad) in a Bio-Rad CFX96 Real-Time PCR System using primers to *GFP* (5'-TACGGCAAGCTGACCCTGAAGT-3' and 5'-CGCACCATCTTCTTCAAGGACG-3') and *ribosomal protein 49* (*rp49*) (5'-TACAGGCCCAAGATCGTGAA-3' and 5'-GCACTCTGTTGTCGATACCC-3'). For each sample, mRNA quantity was determined using the standard curve for each gene analyzed. To determine the relative levels of GFP mRNA over a diurnal cycle, *GFP* mRNA levels were divided by *rp49* mRNA levels for each time point and plotted as the *GFP*/*rp49* mRNA ratio. To quantify *GFP* mRNA in different *tim*-CEG strains at the wild-type (E1-E2) peak, *GFP*/*rp49* values were normalized to the E1-E2 value at ZT14.

II.D. Western blotting

Flies were entrained in a 12-h light/12-h dark (LD) incubator for at least 3 days, collected at the indicated time points, and frozen. Isolated frozen fly heads were homogenized in radioimmunoprecipitation assay (RIPA) buffer (20 mM Tris at pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.05 mM EGTA, 10% glycerol, 1% Triton X-100, 0.4% sodium deoxycholate) containing 0.5 mM PMSF (phenylmethylsulfonyl fluoride), 10 μg/ml aprotinin, 10 μg/ml leupeptin, 2 μg/ml pepstatin A, 1 mM Na_3VO_4, and 1 mM

NaF. This homogenate was sonicated 3 to 5 times for 10 s each time, using a Misonix XL2000 model sonicator at a setting of 3 and then centrifuged at 20,000 g for 10 min. The supernatant was collected as RIPA S extract, and protein concentration was determined by the Bradford assay. Equal amounts of RIPA S extract were run, transferred, and probed with guinea pig anti-CWO (GP-27), 1:5,000; guinea pig anti-CLK (GP-50), 1:5,000; guinea pig anti-PER (GP-73), 1:5,000; rabbit anti-HA (Abcom) and mouse anti-beta-actin (Abcom), 1:20,000. Horseradish peroxidase-conjugated secondary antibodies (Sigma) against guinea pig, rabbit and mouse were diluted 1:5,000. Immunoblots were visualized using ECL plus (GE) reagent.

II.E. ChIP-seq and RNA-seq libraries preparation

II.E.1. ChIP-seq library preparation

ChIP DNA was prepared following II.A. DNA sequencing library construction was performed using NEBNext® Ultra™ DNA Library Prep Kit for Illumina (E7370S) following manufacturer's instructions for end repair, adaptor ligation and size selection. The DNA products were then used as template for PCR amplification for 12 cycles following the PCR condition in the manufacturer's instruction and after purification, the eluted DNA targets were send out for sequencing.

II.E.2. RNA-seq library preparation

RNA extraction was performed following II.C. RNA library construction was performed using NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina (E7420S)

following manufacturer's instructions. The produced cDNA products were then used as template for PCR amplification for 12 cycles following the PCR condition in the manufacturer's instruction and after purification, the eluted DNA targets were then send out for sequencing.

II.F. Bioinformatic analyses for ChIP-seq and RNA-seq data

II.F.1. ChIP-seq mapping and peak finding

Immunoprecipitated DNA libraries were sequenced using Illumina Hiseq 2500 system. Sequences from the different libraries (fastq format) were first mapped to the *Drosophila* genome (version dm3) using bowtie2. Only those reads that mapped uniquely to the *Drosophila* genome were sorted using the samtools suite, and used for further analysis (Menet et al., 2012, 2014). Peak calling was performed using program called findPeaks from HOMER software suite (http://homer.ucsd.edu). Briefly, findPeaks loads tags from each chromosome, adjusting them to the center of their fragments, or by half of the estimated fragment length in the 3' direction. It then scans the entire genome looking for fixed width clusters with the highest density of tags. As clusters are found, the regions immediately adjacent are excluded to ensure there are no "piggyback peaks" feed off the signal of large peaks. By default, peaks must be greater than 2x the peak width apart from one another. This continues until all tags have been assigned to clusters. Significant peaks were computationally assigned to a gene. The following criteria were used: FDR rate threshold = 0.001, p-value over local region required = 1.00e-04, fold over local

region required = 4.00. Visualization of the ChIP-seq signal was performed using the bw file (from the MACS analysis) and the IGV software.

II.F.2. RNA-seq data analysis

RNA (cDNA) libraries were sequenced using Illumina Hiseq 2500 system. Sequences from the different libraries (fastq format) were first mapped to the *Drosophila* genome using with tophat. Uniquely mapped sequences from the tophat output files (bam format) were then used for further analysis. Cufflinks was used for the analysis of gene expression. GO analysis was performed using PANTHER classification system (Mi et al., 2017)

II.G. *Drosophila* activity monitoring and behavior analysis

One to three day old male flies were entrained for three days in LD and transferred to DD for seven days at 25°C. Locomotor activity was monitored using the *Drosophila* Activity Monitor (DAM) system (Trikinetics). Analyses of period, power and rhythm strength during DD was carried out using ClockLab (Actimetrics) software as previously described (Pfeiffenberger et al., 2010).

CHAPTER III

CWO ENHANCES PER MEDIATED RHYTHMS IN TRANSCRIPTIONAL REPRESSION BY COMPETING WITH CLK IN E-BOX BINDING*

III.A. Introduction

Circadian clocks in eukaryotes keep time via cell-autonomous transcriptional feedback loops. A well-characterized example of such a transcriptional feedback loop is in *Drosophila*, where CLK-CYC complexes activate transcription of *per* and *tim*, increasing levels of PER-TIM complexes feedback to repress CLK-CYC activity, and degradation of PER and TIM permits the next cycle of CLK-CYC transcription. The timing of transcriptional events in this feedback loop coincide with, and are controlled by, rhythms in CLK-CYC binding to E-boxes. PER was previously found to inhibit CLK-CYC binding to E-boxes *in vitro* (Lee et al., 1999), which suggests that the rhythmic transcription of CLK target genes are mediated by PER-dependent rhythms in E-box binding by CLK-CYC. However, the mechanism by which CLK-CYC heterodimers are removed from E-boxes by PER during repression is not well understood. Interestingly, co-expression of another transcription factor CWO strongly enhances PER-mediated repression in cultured S2 cells (Kadener et al., 2007), suggesting that PER is unable to efficiently remove CLK from DNA in the absence of other transcription repressors.

Previous studies demonstrated that CWO, a bHLH-ORANGE transcriptional factor (Davis and Turner, 2001), is a direct target of CLK-CYC (Abruzzi et al., 2011; Kadener et al., 2007; Matsumoto et al., 2007). In *Drosophila* S2 cells, overexpression of CWO reduces the basal transcription of *per*, *tim*, *vri* and *Pdp1ε* promoter-driven luciferase reporter genes (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007). Furthermore, in the presence of PER, CWO represses CLK mediated transcription 5-10 fold in S2 cells, indicating that CWO is a strong transcription repressor that can cooperates with PER to repress CLK-CYC-mediated transcription (Kadener et al., 2007; Matsumoto et al., 2007). In *cwo* mutants or *cwo* RNAi knockdown flies, the levels of *per*, *tim*, *vri* and *Pdp1ε* mRNAs are increased during the early to mid-morning (Kadener et al., 2007; Matsumoto et al., 2007). These results suggest that CWO co-represses CLK-CYC activity along with PER during the end of a cycle (Kadener et al., 2007; Matsumoto et al., 2007). However, the mechanism through which CWO represses CLK-CYC-mediated gene transcription remains unknown.

In this chapter I demonstrate that CWO and CLK bind core clock gene E-boxes in a reciprocal pattern across the circadian cycle *in vivo*, which suggests that CWO competes with CLK to bind E-boxes. I also show that CWO acts to decrease CLK binding to *tim* E-boxes during early morning, when PER binds CLK-CYC to reduce its binding to DNA (Yu et al., 2006), but not during early night when CLK-CYC strongly bind E-boxes in the absence of PER. These results suggest a model for CWO function where CWO has low DNA binding affinity compared to CLK-CYC complexes during the activation phase, but has higher affinity compared to CLK-CYC-PER complexes,

and is thus capable of removing CLK-CYC-PER complexes from E-boxes to consolidate and maintain repression. Constitutive CWO binding to the *tim* promoter in Clk^{out} flies and little or no CWO binding in per^{01} flies supports our model for CWO repression. As a whole, these results suggest that CWO co-represses CLK-CYC activity with PER by competing with CLK-CYC-PER complexes for E-box binding, therefore promoting the transition to off-DNA repression.

III.B. Results

III.B.1. CWO is present at constant levels and rhythmically binds E-boxes in a reciprocal pattern compared to CLK

Earlier studies demonstrated that *cwo* mRNA cycles in phase with *per*, *tim*, *vri*, and *Pdp1*, but with a higher basal level, and thus lower amplitude (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). To determine whether CWO protein levels also cycle, western analysis was carried out using head extracts from wild-type flies collected every 4 hours in an LD cycle. I find that the levels of CWO do not change throughout an LD cycle (Fig III.1). Given that *cwo* mRNA levels cycle, it is possible that constant CWO levels result from post-transcriptional regulation or a long half-life.

CWO contains a bHLH domain, a structural motif that characterizes a family of E-box binding transcription factors (Chen et al., 2014; Dawson et al., 1995; Littlewood and Evan, 1995; Massari and Murre, 2000), which suggests that CWO may regulate CLK-CYC target gene transcription via E-box binding. Previous ChIP-on-chip and gel-

shift analyses in S2 cells demonstrated that CWO specifically binds to the E-box of core clock genes (Lim et al., 2007; Matsumoto et al., 2007), however it is still unknown whether CWO binds those core clock genes *in vivo*, and whether the binding intensity changes throughout the day. To test these possibilities, ChIP assays were carried out on wild-type flies collected in the early morning (ZT2) and in the early night (ZT14) using CWO and CLK antisera. Fragments containing upstream E-boxes from *tim*, *per*, *Pdp1* and *vri*, which are necessary for high-amplitude mRNA cycling *in vitro* or *in vivo* (Blau and Young, 1999; Cyran et al., 2003; Darlington et al., 2000; Hao et al., 1997, 1999; Lyons et al., 2000; McDonald et al., 2001; Moore and Eichler, 1972), were amplified from the immunoprecipitates and then quantified. In CWO immunoprecipitates, the *tim*, *vri* and *Pdp1* E-box containing fragments were two to threefold more abundant at ZT2 than at ZT14 (Fig III.2A), suggesting that CWO binding is time-dependent, though the dynamic binding of CWO on the *per* E-box fragment is less robust than the others. Importantly, this temporal binding pattern is antiphase to CLK binding, as CLK shows high binding intensity during the night at ZT14 and low binding during the daytime at ZT2 (Fig III.2B), consistent with previous results (Abruzzi et al., 2011; Yu et al., 2006).

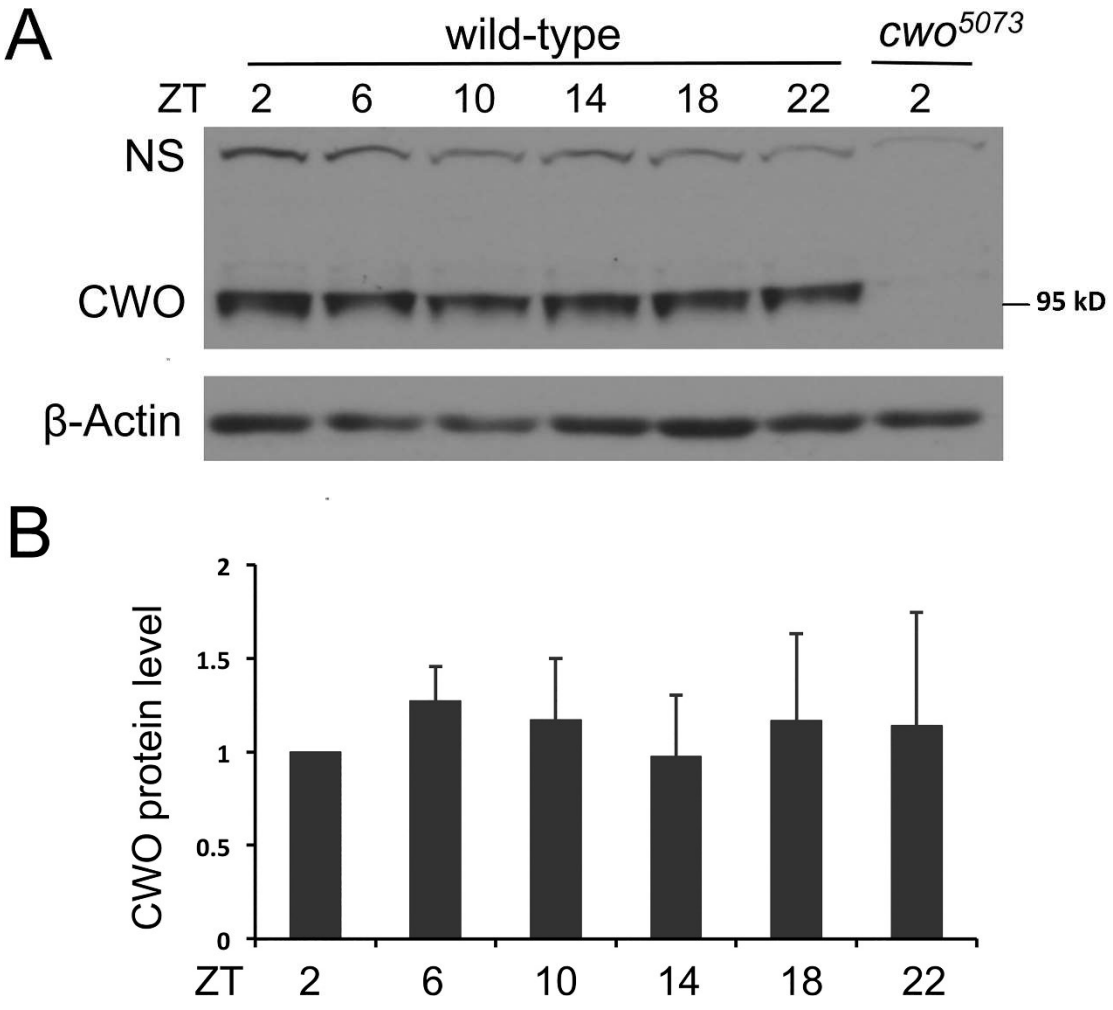


Fig III.1. CWO protein is present at constant levels in fly heads.

(A) Western blot of head extracts from wild-type and *cwo*^{5073} flies collected at the indicated times were probed with CWO antiserum. β-Actin or a nonspecific band (NS) were used as loading controls. **(B)** Quantification of CWO levels in the blot from panel A and two additional western blots containing samples from independent collections. Error bars indicate the SD (n = 3).

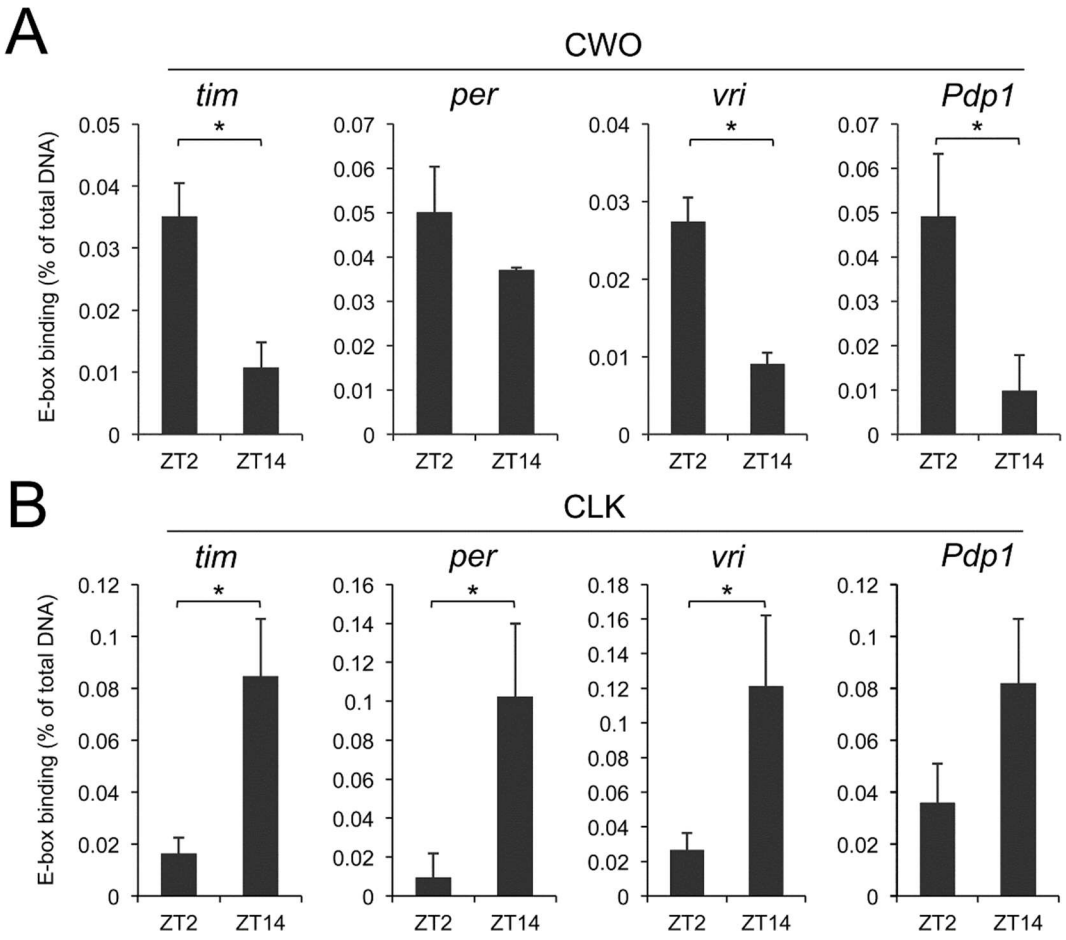


Fig III.2. CWO rhythmically binds E-boxes of core clock genes in antiphase to CLK.

(A) ChIP assays were performed on wild-type flies collected at ZT2 and ZT14. The relative level of CWO binding to *tim*, *per*, *vri* and *Pdp1* E-boxes was determined by qPCR analysis of samples immunoprecipitated with CWO antiserum (see Materials and Methods). The mean values of three independent ChIP assays were calculated and plotted. Error bars indicate the SEM (n = 3, *p<0.05, Student's t-test). **(B)** ChIP assays of samples immunoprecipitated with CLK antiserum were performed, quantified and plotted as described for panel A. Error bars indicate the SEM (n = 3, *p<0.05, Student's t-test).

III.B.2. CWO and CLK bind tandem E-boxes upstream of *tim*

The reciprocal binding pattern of CLK and CWO implies that these transcription factors compete for E-box binding. If so, both CLK and CWO must occupy the same E-boxes. To test this possibility, I determined how mutating E-boxes upstream of *tim* affected CLK and CWO binding. The circadian enhancer upstream of *tim* is comprised of two tandem E-boxes that are spaced seven nucleotides apart (McDonald et al., 2001; Paquet et al., 2008), a structure that is conserved among core clock genes in various species (Rey et al., 2011). Both of these E-boxes were indispensable for *tim* mRNA expression in S2 cells (McDonald et al., 2001), suggesting that these tandem E-box motifs are binding sites for both CLK and CWO. To determine if this is the case, a series of 136bp fragments from the *tim* promoter containing an E-box1 (E1) mutant (mE1-E2), an E-box 2 (E2) mutant (E1-mE2), an E1 and E2 double mutant (mE1-mE2) or a control with wild-type E-boxes (E1-E2) were generated, inserted into the pHPdestGFP vector (Boy et al., 2010), and targeted to the attP18 genomic site (Fig III.3A).

To confirm that this promoter fragment is sufficient to drive rhythmic expression, I carried out quantitative reverse transcription-PCR (qRT-PCR) to monitor *GFP* mRNA levels in flies collected every 4-h during an LD cycle. Quantification of *GFP* mRNA levels in flies with WT *tim* promoter shows a ~10-fold diurnal rhythm with a peak at ZT14 and a trough at ZT2 to ZT6 (Fig III.S1A), consistent with timing and amplitude of *per* and *tim* mRNA cycling in wild-type flies (Hardin et al., 1990; Sehgal et al., 1995). However, even at the normal *tim* mRNA peak (ZT14), mE1-E2, E1-mE2 and mE1-mE2 flies express little or no eGFP mRNA (Fig III.S1B), indicating that both E1 and E2 are

indispensable for expression of *tim* mRNA *in vivo*. This result is consistent with previous *tim*-luciferase reporter results in S2 cells (McDonald et al., 2001).

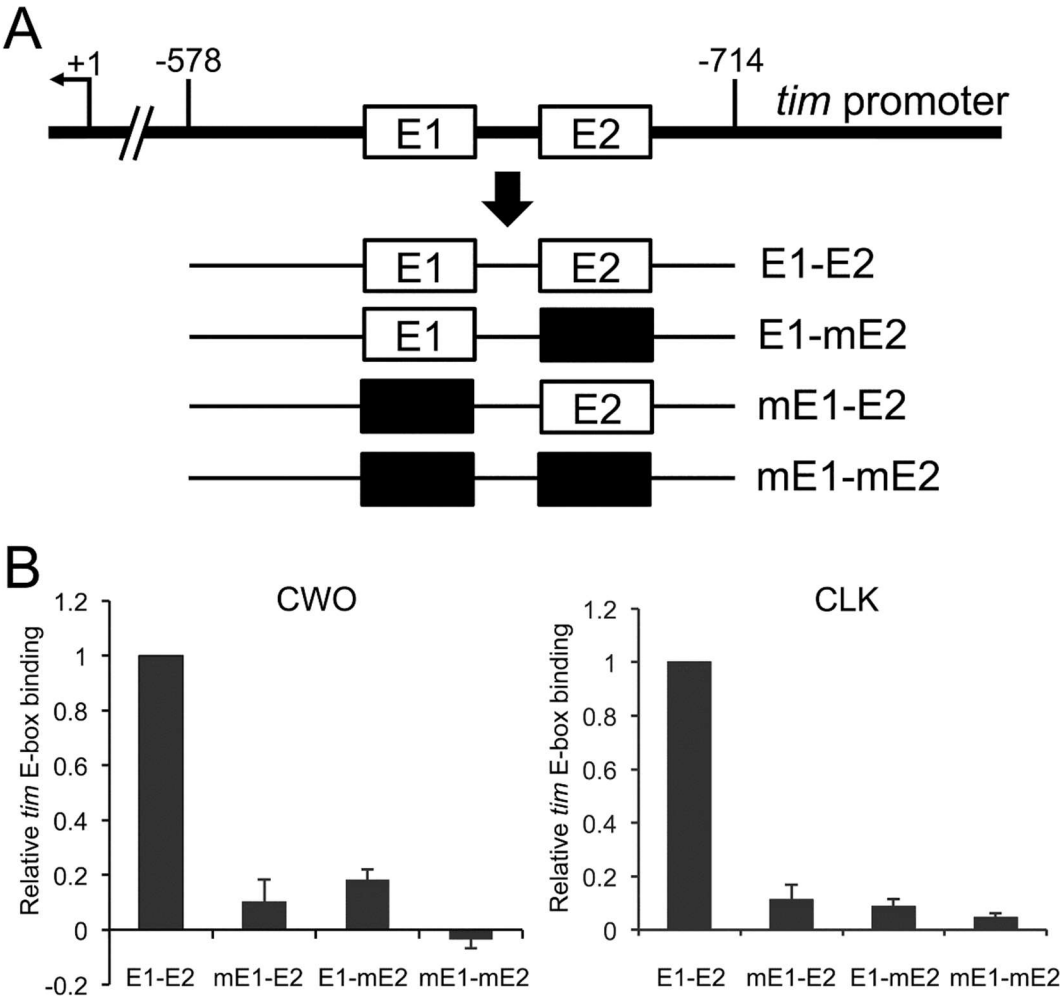


Fig III.3. CWO and CLK bind the same tandem E-boxes in the *tim* circadian enhancer.

(A) Schematic diagram of wild-type and mutant *tim* circadian enhancer transgenes. A 136bp *tim* circadian enhancer fragment that extends from -578 to -714 relative to the *tim* transcription start (+1; see (McDonald et al., 2001)) was used to generate transgenes

with wild-type or mutant combinations of tandem E-box 1 (E1) and E-box 2 (E2) motifs. Wild-type and mutant *tim* circadian enhancer fragments were generated via PCR, cloned into the pHPdesteGFP reporter plasmid, and used to generate transgenic flies via PhiC31 recombination (see Materials and Methods). The resulting *tim* circadian enhancer transgenes contain wild-type E1 and E2 (E1-E2), mutant E1 and wild-type E2 (mE1-E2), mutant E2 and wild-type E1 (E1-mE2), or mutant E1 and E2 (mE1-mE2) E-boxes. Black boxes, mutant E1 or E2 E-boxes; double backslash, virtual break in the *tim* promoter sequence. **(B)** ChIP assays on flies containing the E1-E2, mE1-E2, E1-mE2 or mE1-mE2 *tim* circadian enhancer transgenes. The relative level of CWO binding at ZT2 or CLK binding at ZT14 was determined as described in Fig III.2A. Relative binding of CWO or CLK was normalized to the maximum E1-E2 value of 1.0, then the means of each data set were calculated and plotted. Error bars represent the SEM (n = 3).

I next carried out ChIP assays using CWO and CLK antisera on the same fly strains to test whether E1 and E2 are required for CWO and CLK binding. At ZT2, when CWO strongly binds to the *tim* promoter, CWO binding intensity was drastically reduced in mE1-E2, E1-mE2 and mE1-mE2 flies compared to WT (Fig III.3B). Likewise, CLK binding intensity was drastically reduced in mE1-E2, E1-mE2 and mE1-mE2 flies compared to WT at ZT14, when CLK binding is strongest (Fig III.3B). These results indicate that both E1 and E2 are indispensable for both CWO and CLK binding to the *tim* circadian enhancer. Given that CWO specifically targets E-boxes in S2 cells by Gel-shift analyses (Lim et al., 2007), I conclude that both CLK and CWO only bind to intact tandem E1-E2 motifs *in vivo*. In mice, CLK-BMAL1 dimers cooperatively bind tandem

E-boxes *in vitro* (Rey et al., 2011; Shimomura et al., 2013), and this may be the case for CWO given the requirement for both E1 and E2 E-boxes.

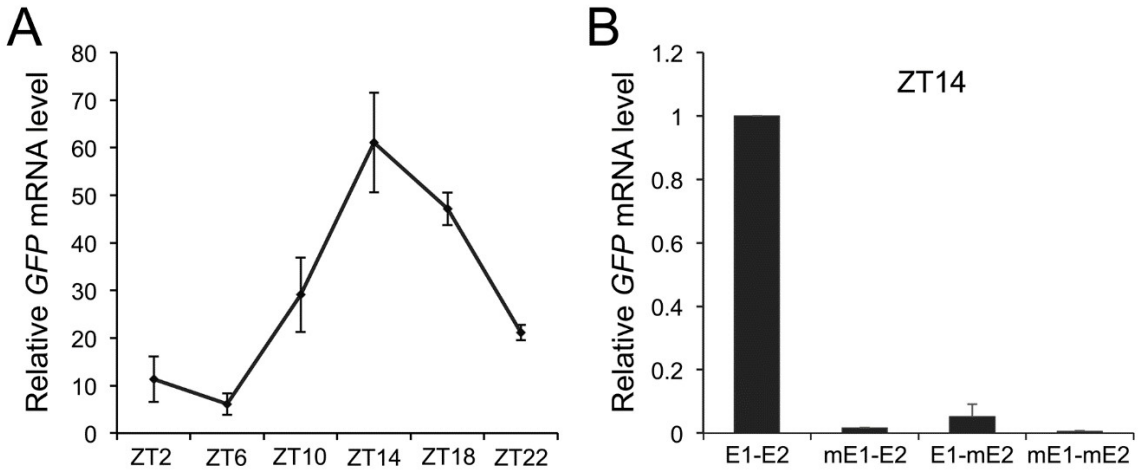


Fig III.S1. *tim* promoter fragments bearing E-box mutations abolish mRNA cycling.

(A) Quantitative PCR (qPCR) was performed to measure GFP mRNA levels in E1-E2 *tim* circadian enhancer flies collected at the indicated times in LD. Relative GFP mRNA values were generated by dividing the GFP mRNA signal by that of ribosomal protein 49 (RP49), which is expressed at constant levels. Error bars represent the SEM (n = 3). (B) Quantification of GFP mRNA levels in E1-E2, mE1-E2, E1-mE2 and mE1-mE2 *tim* circadian enhancer transgenic flies collected at ZT14 as described in panel A. Relative GFP mRNA levels were normalized to the E1-E2 *tim* circadian enhancer fly value, which was designated as 1.0, then the means of each data set were calculated and plotted. Error bars represent the SEM (n = 3).

III.B.3. CWO represses CLK binding to tim promoter during transcription repression

Previous studies showed that increasing the level of CWO expression reduces *per*, *tim*, *vri* and *Pdp1ε* mRNA levels in S2 cells and that their trough mRNA levels are higher in *cwo* mutant or knockdown flies, indicating that CWO acts to repress CLK-mediated gene transcription *in vitro* and *in vivo* (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). Given that CWO and CLK bind to the same E-box motif, I wondered whether CWO represses CLK-mediated transcription by inhibiting CLK binding. To test this possibility, ChIP assays were carried out using CLK antiserum on wild-type and cwo^{5703} flies at the trough (ZT2) and peak (ZT14) times of CLK-CYC target gene transcription and mRNA abundance in LD. Although cwo^{5703} mutants lengthen the period of activity rhythms by 2–3h in DD (Kadener et al., 2007; Lim et al., 2007), the peak and trough phases of CLK-CYC target gene transcription and mRNA abundance are comparable in cwo^{5703} mutants and wild-type flies in LD (Kadener et al., 2007; Lim et al., 2007). I find that CLK binds *tim* E-boxes with a robust rhythm in wild-type flies and a lower amplitude rhythm in the cwo^{5703} mutant (Fig III.4A). However, the intensity of CLK binding in cwo^{5703} is significantly increased at ZT2 compared to wild-type, indicating that CWO acts to reduce CLK-CYC binding at the trough of its binding cycle (Fig III.4B). Given that CWO strongly binds *tim* E-boxes at ZT2 (Fig III.2A), we propose that CWO inhibits CLK-CYC binding during the repression phase by antagonizing PER-CLK-CYC complexes to maintain off-DNA repression. There was no significant difference in CLK binding between cwo^{5703} and wild-type at ZT14 (Fig

III.4B), despite decreased peak levels of *per*, *tim*, *vri* and *Pdp1* mRNA at ZT14 in *cwo* mutant and RNAi knockdown flies (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008), suggesting that CWO has little impact on CLK-CYC binding in the absence of PER.

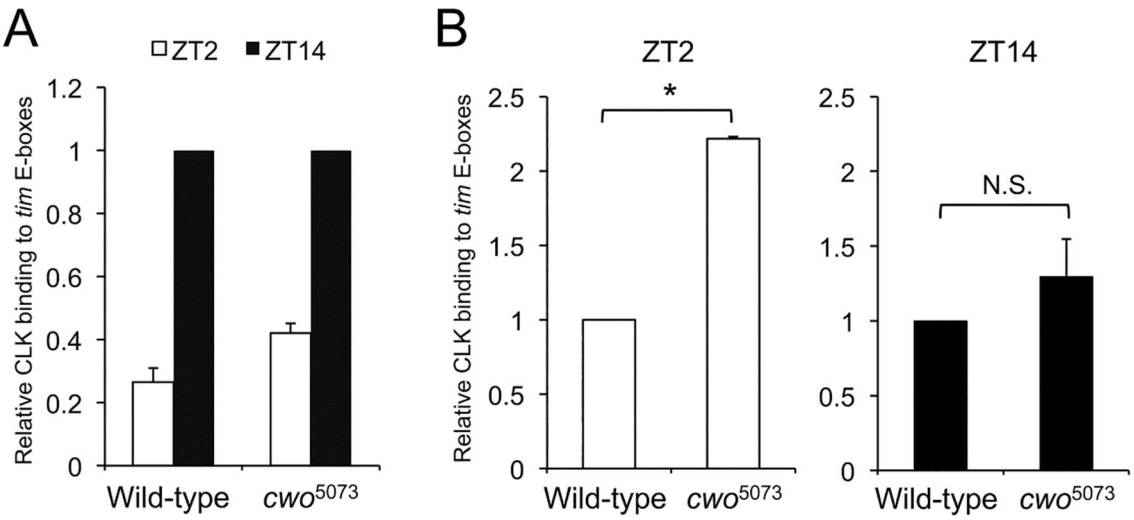


Fig III.4. CWO reduces CLK binding to *tim* E-boxes during transcriptional repression.

ChIP assays were performed on wild-type and *cwo*^{5073} flies collected at ZT2 and ZT14, and the relative level of CLK binding to *tim* E-box-containing fragments was determined as described in Fig III.2A. **(A)** CLK binding signal was normalized to a ZT14 value of 1.0 for wild-type and *cwo*^{5073} flies, respectively, and the mean values at ZT2 from each data set (n = 3) were calculated and plotted. **(B)** CLK binding signal from *cwo*^{5073} flies was normalized to a wild-type value of 1.0 at ZT2 and ZT14, respectively, and the mean values of each data set were calculated and plotted. Error bars represent the SEM (n = 3, *p<0.05, Student's t-test).

III.B.4. PER is required for CWO to displace CLK-CYC binding on E-boxes

Given that CWO represses CLK binding at ZT2 in the early morning but not at ZT14 during the night (Fig III.4B), it is possible that PER is necessary for CWO to compete with CLK for E-box binding since PER accumulates to high levels in the nucleus around dawn and is at low levels in the cytoplasm around dusk (Curtin et al., 1995). This possibility is supported by an earlier study showing that CWO and PER work synergistically to repress CLK-mediated gene expression in S2 cells (Kadener et al., 2007). I therefore hypothesize that CWO competes with CLK-CYC heterodimers for E-box binding only when PER binds CLK-CYC, thereby reducing their affinity for E-box binding. To determine the impact of CLK and PER on CWO binding to E-boxes, I performed ChIP assays using CWO antiserum on wild-type, Clk^{out} and per^{01} flies collected at ZT2 and ZT14 in LD. In Clk^{out} flies, which necessarily lack CLK-CYC heterodimers (Mahesh et al., 2014), CWO is bound to *tim* E-boxes at both ZT2 and ZT14 with binding signals comparable to high CWO binding in wild-type flies at ZT14 (Fig III.5A). In contrast, in per^{01} flies, which lack PER-dependent repression of CLK-CYC activation (Darlington et al., 1998), low binding signals of CWO were detected at ZT2 and ZT14, indicating that PER is indeed required for CWO to bind E-boxes (Fig III.5A). Moreover, CWO binding was significantly increased in Clk^{out} versus wild-type flies at ZT14, indicating that CLK-CYC binding at ZT14 reduces CWO binding. Likewise, a significant increase in CWO binding was also seen in wild-type versus per^{01} flies at ZT2, indicating that PER enhances CWO binding (Fig III.5A).

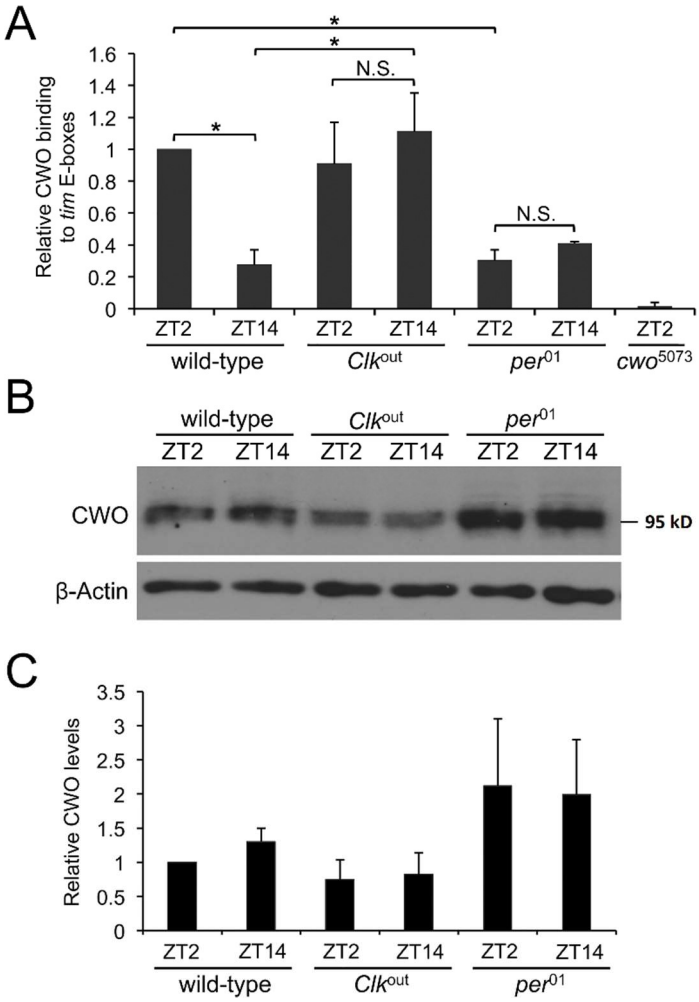


Fig III.5. PER is required for CWO to compete with CLK-CYC for E-box binding. **(A)** ChIP assays were performed on wild-type, *Clk*^{out} and *per*^{01} flies collected at ZT2 and ZT14, and *cwo*^{5073} flies collected at ZT2, as described in Fig III.2A. The relative level of CWO binding to *tim* E-box-containing fragments was determined as described in Fig III.2A. The signal from each sample was normalized to the wild-type ZT2 value of 1.0, then the means of each data set (n = 3) were calculated and plotted. Error bars represent the SEM (n = 3, *significantly different, ANOVA followed by Student-Newman-Keuls post-hoc test). **(B)** Western blot of head extracts from the same genotypes shown in panel A were probed with CWO antiserum. **(C)** Quantification of CWO levels in the blot from panel B and two additional Western blots containing samples from independent collections. Error bars indicate the SD.

To determine whether differences in CWO binding in *Clk*^{out} and *per*^{01} flies were due to differences in CWO protein levels, I carried out western analysis using head extracts from these mutants collected at ZT2 and ZT14. Since *cwo* transcription is regulated in part by the transcriptional feedback loop, CWO protein levels are slightly lower in *Clk*^{out} flies and slightly higher in *per*^{01} flies (Fig III.5B, C). However, the lower levels of CWO in *Clk*^{out} resulted in higher E-box binding, and higher CWO protein levels in *per*^{01} resulted in lower E-box binding. This result indicates that the differences in CWO-E-box binding are not due to altered CWO protein levels, but due to the relative DNA binding affinities of CWO and CLK in these mutants. These results, taken together, strongly support and extend the model described by Kadener et al., 2007, for CWO binding as it relates to CLK-CYC repression. When CLK-CYC targets are activated, CLK-CYC bind DNA with higher affinity than CWO, thus CLK binding is not altered in the presence or absence of CWO. When CLK-CYC targets are repressed, PER binds CLK-CYC complexes and decreases its DNA binding affinity, CWO is then able to compete with CLK-CYC-PER for E-boxes binding, thereby enhancing PER mediated removal of CLK-CYC-PER complexes from the DNA (Fig III.6). Although we can't exclude the possibility that PER enables CWO E-box binding independent of its interaction with CLK-CYC, the available evidence strongly supports the model proposed.

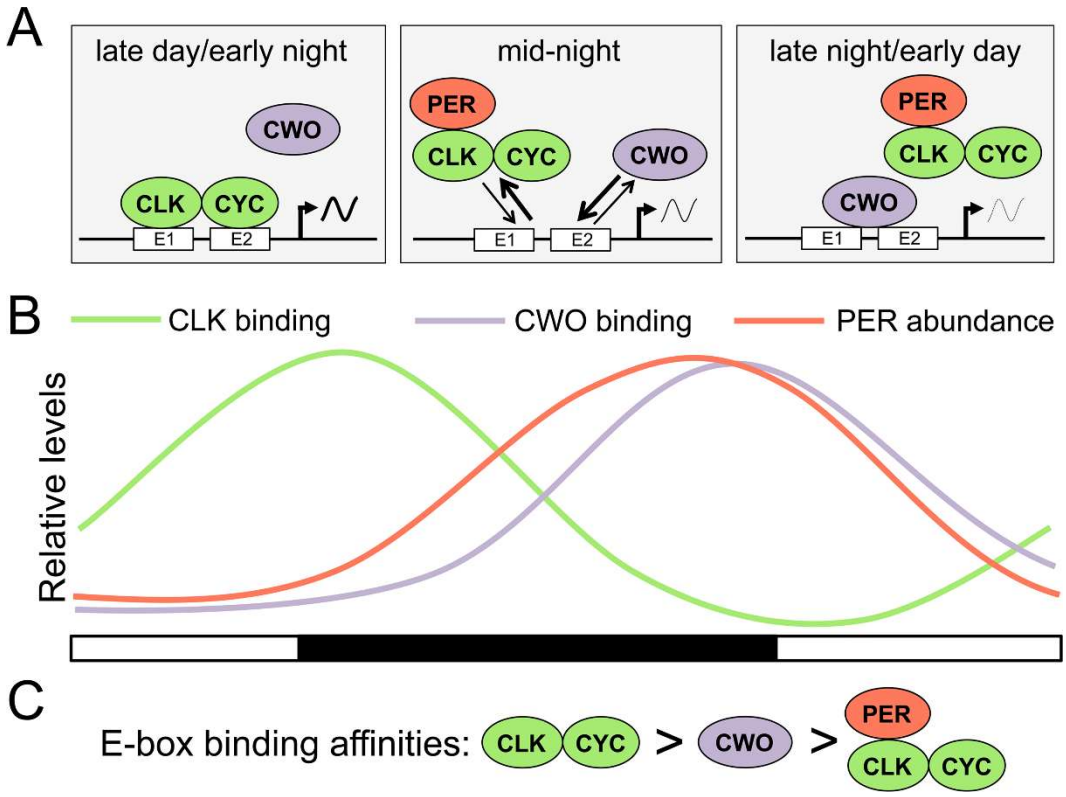


Fig III.6. Model for PER dependent binding competition between CWO and CLK-CYC on E-boxes.

(A) Diagrams depicting clock protein interactions and E-box binding at different times of day (gray rectangles). During late day/early night, CLK-CYC (green ovals) initiate the transcription cycle by binding to E-boxes (white rectangle) in the presence of CWO (purple oval), which is unable to bind E-boxes. During mid-night, PER (red oval) enters the nucleus and interacts with CLK-CYC, producing PER-CLK-CYC complexes that allow CWO binding to E-boxes. During late night/early day, the high PER levels insure efficient PER-CLK-CYC complex formation, thus allowing strong CWO binding to E-boxes. **(B)** Graph showing the relative levels of CLK-CYC binding (green line), CWO binding (purple line) and PER abundance (red line) during a light (white rectangles) and dark (black rectangles) cycles. **(C)** Proposed E-box binding affinities, where CLK-CYC binding is greater than (>) CWO, and CWO binding is greater than PER-CLK-CYC.

III.C. Summary

In this chapter I demonstrate that the transcription factor CWO competes with CLK-CYC in E-box binding, thus enhancing the removal of CLK-CYC from E-boxes to maintain transcriptional repression. This process is PER dependent, which suggests that PER and CWO cooperate to maintain a transcriptionally repressed state by removing CLK-CYC from E-boxes. These results demonstrate that PER-TIM require CWO to effectively repress circadian transcription. Given that CWO orthologs DEC1 and DEC2 also target E-boxes bound by CLOCK-BMAL1, this mechanism may function to repress transcription in other animals including humans.

CHAPTER IV

GENOME-WIDE ANALYSIS OF CWO'S FUNCTION IN TRANSCRIPTIONAL REGULATION

IV.A. Introduction

The circadian feedback loops that control the daily rhythms in physiology, metabolism and behavior are primarily regulated by rhythmic binding of CLK-CYC to E-boxes of the core circadian oscillator genes in *Drosophila*. Results from Chapter III demonstrate that CWO is a transcriptional repressor and competes with CLK in E-box binding, therefore enhancing transcriptional repression during the early night. A previous ChIP-on-chip assay identified ~1500 CWO binding sites in S2 cell culture (Matsumoto et al., 2007), suggesting that CWO broadly regulates gene expression, thus potentially controlling circadian rhythms and many other biological processes. However, CWO protein was overexpressed in these S2 cell experiments, leading to potential false-positive interactions: CWO may bind to many targets that it would not bind endogenously. Moreover, the binding sites of CWO detected in S2 cells does not fully reflect CWO binding *in vivo,* as the core clock genes *per* and *tim* that have the E-box sequences in their promoter region were not detected in the experiment. These potential false-positives and false-negatives then prompted us to carry out *in vivo* ChIP-seq experiments to identify potential CWO targets and to characterize the binding pattern of CWO at genome-wide level.

In *cwo* mutant or RNAi knockdown flies, mRNA abundance of clock genes is reduced during the transcriptional activation phase, indicating that CWO also acts to promote transcription (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). However, the mechanism of how CWO promotes transcription is unknown. Data from Chapter III show that CLK binding is not affected by CWO during the activation phase (Fig III.4B), suggesting that CWO may indirectly regulate transcription activation during the early morning, independent of its role during the night in directly competing with CLK for E-boxes binding. Therefore, identifying genes that are differentially expressed in *cwo* mutant strain could provide clues in understanding the mechanism of how CWO indirectly acts as activator to promote gene transcription.

In this chapter I demonstrate that CWO binds ~500 genes in the genome by ChIP-seq analysis. I also find a substantial overlap between CWO and CLK target genes, suggesting that CWO plays a potential role in regulating CLK-mediated transcription globally. Motif analysis indicates that the canonical E-box sequences are highly enriched among CWO target regions, which coincides with previous observations that CWO specifically binds E-boxes in S2 cells and *in vivo* (Lim et al., 2007; Matsumoto et al., 2007; Zhou et al., 2016). Using RNA-seq I identify 588 upregulated genes and 582 downregulated genes in the cwo^{5073} mutant, with few CWO direct target in common, suggesting that CWO mainly regulates gene expression in an indirect manner. Interestingly, CLK is found hyper-phosphorylated in cwo^{5073} mutants during the activation phase, potentially due to the differential expression of a subset of genes encoding kinases and phosphatases. Taken together, my study suggests that CWO

indirectly regulates CLK phosphorylation by changing the expression pattern of genes encoding kinases or phosphatases, thus promoting transcription of clock genes.

IV.B. Results

IV.B.1. HA tagged CWO protein is detectable by anti-HA and can partially rescue *cwo* mutant phenotype

To detect and immunoprecipitate CWO with high sensitivity and specificity, a BAC transgene that expresses C-terminal HA-tagged CWO was generated (see the Materials and Methods). As expected, HA antibody specifically detects CWO-HA on western blots (Fig IV.1A). To determine whether CWO-HA protein expresses in the same pattern as endogenous CWO protein, western analysis was carried out using head extracts from *cwo-HA; cwo*^{5073} transgenic flies collected every 4 hours in LD. CWO-HA levels do not change throughout an LD cycle (Fig IV.1B), similar to previous western analysis showing the temporal expression pattern of endogenous CWO protein (Fig III.1). Behavior assays were then performed to determine whether CWO-HA can rescue the long period behavior phenotype observed in *cwo*^{5073} flies. I find that *cwo-HA; cwo*^{5073} flies had a period of ~24.6 h in DD (Table IV.1), which is ~1.5h shorter than in *cwo*^{5073} flies, but ~1h longer than wild-type (w^{1118}) flies. These results demonstrate that the CWO-HA fusion protein is functional, and will be a good tool for ChIP-seq analysis.

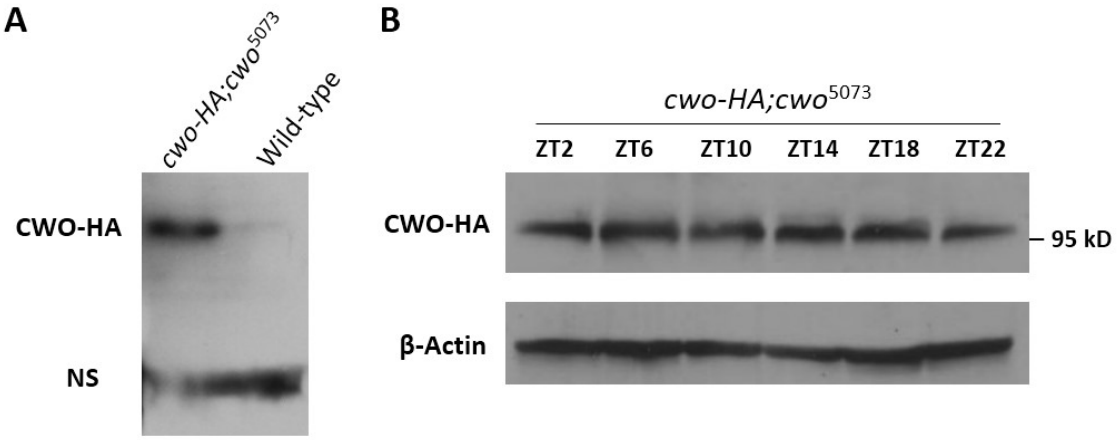


Fig IV.1. CWO-HA protein is present at constant levels in *cwo-HA; cwo*^{5073} transgenic fly heads.

(A) Western blot of head extracts from *cwo-HA; cwo*^{5073} and wild-type flies probed with anti-HA. **(B)** Western blot of head extracts from *cwo-HA; cwo*^{5073} collected at the indicated times were probed with anti-HA. β-Actin or a nonspecific band (NS) were used as loading controls.

Genotype	Total (n)	% Rhythmic	Period (mean ± SEM)
Wild-type	11	100 (11)	23.50 ± 0
cwo^{5073}	8	85.7 (7)	26.07 ± 0.53
cwo-HA; cwo^{5073}	32	84.37 (27)	24.64 ± 0.05

Table IV.1. *cwo-HA* partially rescues *cwo*^{5073} mutant phenotype.

IV.B.2. Identification of CWO direct target genes in *Drosophila*

To identify CWO direct target genes, HA antibody was used to perform ChIP-seq on *cwo-HA; cwo*^{5073} head extracts at two time points (ZT2 and ZT14) in a LD cycle (see the Materials and Methods). Total of 393 and 549 CWO-bound DNA fragments were identified from samples collected at ZT2 and ZT14, respectively, using the HOMER software suite (see the Materials and Methods). Notably, CWO binding peaks are preferentially found in the promoter regions, with 53.2 % and 42.4 % of peaks within Promoter-transcription start sites (TSS) (defined from -1kb to +100bp) from samples collected at ZT2 and ZT14, respectively (Table.IV.2). These data reveal a preference for CWO binding in the proximal promoter regions of potential CWO target genes at both time points. To further identify CWO target genes at both time points, peaks that map to intergenic regions or different regions in the same gene were deleted, which results in 325 and 437 direct target genes at ZT2 and ZT14, respectively, with 270 genes in common (Fig IV.2A). CWO target genes at ZT2 and ZT14 were then combined, since genes that appear at least at one time point were regarded as CWO targets, which results in a list of 492 target genes. Importantly, the known CWO direct targets, *per*, *tim*, *vri* and *Pdp1* rank in the top third of the 492 target genes (Fig IV.2B, Fig IV.3). A previous CWO ChIP-on-chip assay carried out in S2 cell culture identified 1103 CWO target genes (Matsumoto et al., 2007). I compared my ChIP-seq data with this data, and found that 154 genes, approximately one-third of all potential direct CWO targets I identified, also appear in the previous microarray results from S2 cells (Fig IV.2C).

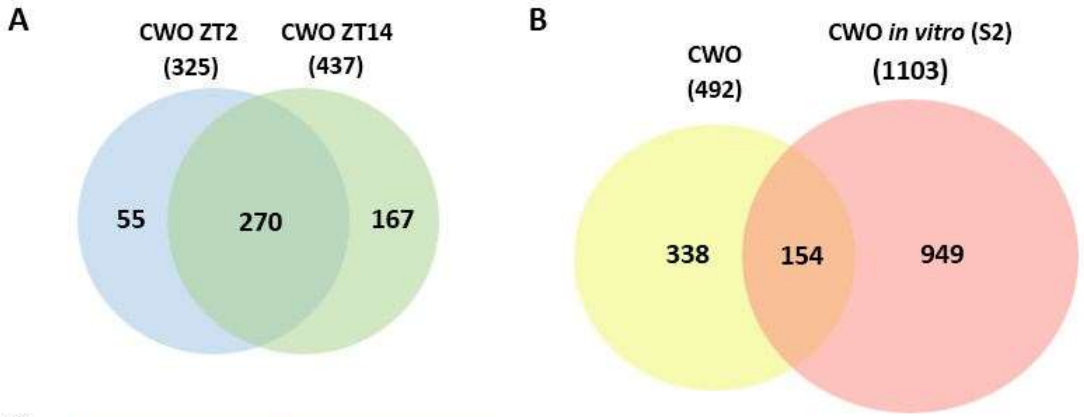
	CWO ZT2	CWO ZT14	CLK ZT14
Total	**393**	**549**	**149**
Promoter-TTS*	209	233	41
3' UTR §	1	1	1
5' UTR §§	38	55	23
Exon	8	21	5
Intron	105	181	60
Intergenic	32	58	19

Table.IV.2. Overall view of the CWO and CLK ChIP-seq peaks in wild-type fly heads.

* Promoter-TTS: defined from -1kb to +100bp of transcription start site

§ 3' UTR: the three prime untranslated region

§§ 5' UTR: the five prime untranslated region



C

Target Genes	Rank
vri	24
per	45
tim	46
pdp1	126

Fig IV.2. ChIP-seq analysis of CWO binding sites and comparison to previous microarray data.

(A) Venn diagrams of CWO ChIP-seq targets at ZT2 (blue) and ZT14 (green). The numbers in the brackets are the total number of targets for each time point. The numbers in the circles or in the overlapped region indicate the numbers of targets that are unique to each category or appear in both category. **(B)** Venn diagrams of comparison between CWO ChIP-seq targets (yellow) and CWO ChIP-on-chip targets (red). The numbers are determined as described in Fig IV.2A. **(C)** Rank of each clock gene on CWO ChIP-seq target list.

Unlike the ChIP-qPCR data reported in Chapter III, where CWO rhythmically binds to CLK target genes with higher binding intensity at ZT2 compared to ZT14 (Fig.III.2), CWO shows a different binding pattern for some clock genes using ChIP-seq

analysis. For *tim*, the binding intensity of CWO is high at ZT2 and low at ZT14; for *per* and *vri*, CWO binding is slightly higher at ZT14 for some regions; for *pdp1*, there is no obvious difference in CWO binding at both time points. (Fig.IV.3). These results indicate that the ChIP-seq data have a bias towards some targets at different time points, possibly caused by PCR duplication during DNA library preparation. The HA tag fused to CWO protein could also interfere with CWO function and DNA binding. Thus, I consider these ChIP-seq data are more of qualitative value (identifying targets) than quantitative value (comparing binding intensities).

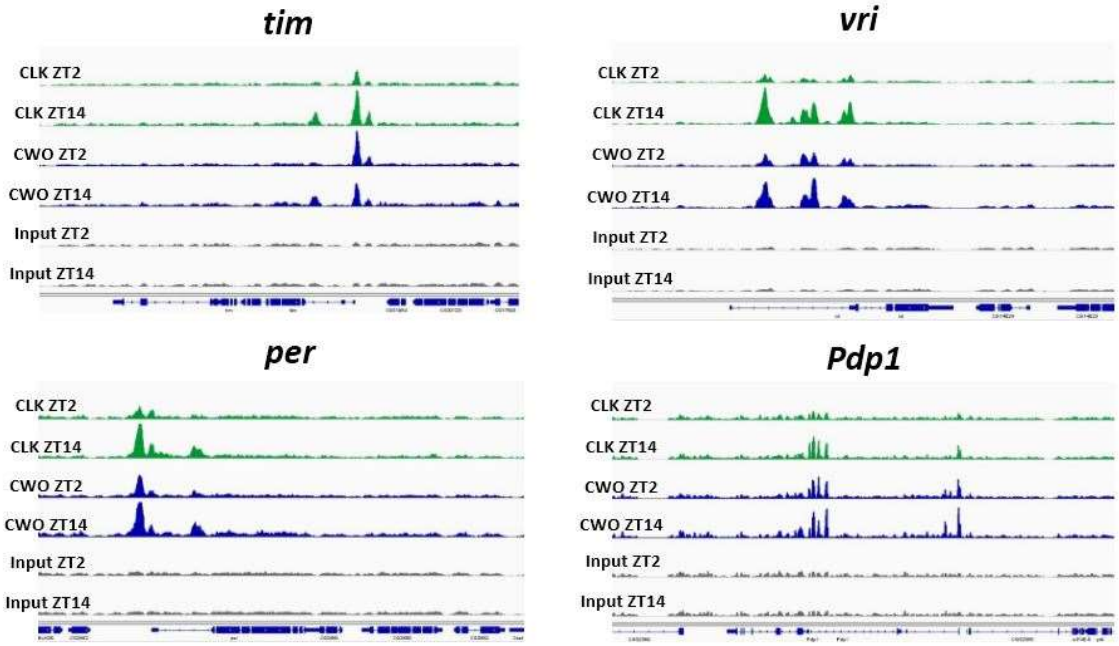


Fig IV.3. Visualizing of CLK and CWO binding to the core clock genes.
ChIP-seq track showing CLK (green) and CWO-HA (blue) binding to the core clock genes *tim*, *vri*, *per* and *Pdp1* at ZT2 and ZT14. Chromatin preps without immunoprecipitation were used as inputs (grey) for the experiment.

Given that CWO specifically binds E-boxes in S2 cells (Lim et al., 2007; Matsumoto et al., 2007) and *in vivo* (Fig.III.3B), motif analyses was then carried out to identify the DNA motif enriched in CWO binding peaks. The CACGTG E-box sequence is the top motif among CWO binding sites by *de novo* motif analysis (Fig IV.4A), and canonical E-box motifs were detected in the 5 top known motifs among the enrichment analysis results (Fig IV.4B), suggesting CWO specifically binds E-box motifs on DNA a genome-wide level.

A

B

CWO

Known motifs	p-value
bHLHE40	1e-153
c-Myc	1e-111
n-Myc	1e-90
Max	1e-88
USF1	1e-82

Fig IV.4. Motif enriched in CWO binding regions

Logos show the *de novo* motif **(A)** and the top five known motifs **(B)** that enriched at the CWO binding peak regions identified by motif analysis.

To explore biological pathways potentially regulated by CWO, genes nearest to CWO binding peaks were assigned to functional groups by GO analysis. GO analysis based on biological process shows 3 statistically significant categories (Table IV.3). Not surprisingly, one of these categories is "behavior", and genes related to circadian rhythms, sleep and locomotor behavior are significantly enriched in this category, which correlates with the function of *cwo* as a biological clock component controlling rhythmic behavior of *Drosophila*. Two other categories are "metabolic process" and "developmental process". Genes related to primary metabolic processes or the regulation of metabolic processes are significantly enriched in the metabolic category; and genes related to tissue development, regulation of cell differentiation and embryo development are significantly enriched in the developmental category. Interestingly, for the developmental category, a sub-category of "central nervous system development", including "synapse organization" and "synapse assembly", is enriched, consistent with the function of *cwo* in dendrite morphogenesis and assembly of the nervous system (Iyer et al., 2013). GO analysis based on molecular function shows two main categories: "binding" and "catalytic activity". It is important to notice that in the "binding" category genes are enriched in nucleic acid binding, sequence-specific DNA binding and transcription factor activity, indicating that a large portion of CWO target genes are transcription factors that potentially bind their targets to regulate gene expression. These downstream transcription factors thus allow *cwo* to indirectly regulate biological processes. All the GO analyses, taken together, suggest that CWO globally binds to its

target genes at genome wide level, directly or indirectly controlling many important biological pathways, such as metabolism, development and behavioral rhythms.

description	frequency	log10 p-value
behavior	**5.96%**	**-4.0115**
rhythmic process	1.20%	-7.3601
circadian rhythm	1.17%	-7.4971
sleep	1.25%	-3.4071
locomotor behavior	1.93%	-3.5709
metabolic process	**52.04%**	**-4.1201**
cellular metabolic process	41.92%	-8.2141
primary metabolic process	44.31%	-4.6066
regulation of metabolic process	17.72%	-11.574
cellular macromolecule metabolic process	31.94%	-2.083
developmental process	**31.18%**	**-10.4345**
cell development	13.87%	-10.6837
embryo development	4.78%	-5.8542
regulation of developmental process	6.52%	-14.5703
regulation of cell differentiation	2.94%	-8.4884
tissue development	10.16%	-11.5717
post-embryonic development	6.56%	-10.9317
central nervous system development	2.43%	-3.4328
synapse organization	2.39%	-4.3314
synapse assembly	1.49%	-3.8722

Table IV.3. GO (biological process) analysis and display of CWO binding peaks. The potential CWO target genes are enriched in three categories: behavior (green), metabolic process (blue) and developmental process (orange) based on biological process analysis.

description	frequency	log10 p-value
binding	**49.13%**	**-14.9368**
nucleic acid binding	3.36%	-11.2044
sequence-specific DNA binding	3.19%	-11.2993
transcription factor activity	3.36%	-11.2044
nucleic acid binding	15.53%	-2.3549
regulatory region nucleic acid binding	1.64%	-6.1483
ribonucleotide binding	8.37%	-3.9032
protein binding	20.27%	-13.194
small molecule binding	10.29%	-3.1903
catalytic activity	**37.74%**	**-2.5566**
ATPase activity	1.07%	-3.2541
phosphoric ester hydrolase activity	1.94%	-2.6005

Table IV.4 GO (functional) analysis and display of CWO binding peaks.
The potential CWO target genes are enriched in two categories: binding (green) and catalytic activity (blue) based on molecular function analysis.

In parallel with CWO ChIP-seq, CLK ChIP-seq assays were also performed for two time points at ZT2 and ZT14. The experimental process was the same as CWO ChIP-seq, except CLK antiserum was used instead of anti-HA. A total of 149 CLK binding peaks representing 113 target genes are detected at ZT14. This number is much less compared to earlier data that reported ~1500 CLK target genes by ChIP-on-chip assays (Abruzzi et al., 2011), probably because I used a different ChIP protocol that employed more stringent wash steps (Zhou et al., 2015). The 22 peaks identified at ZT2

is even lower than that at ZT14, primarily because CLK binding at ZT2 is so low that binding signals are not detected by the peak calling program. In the previous CLK ChIP-on-chip assay, 500 genes were mapped with cycling of CLK-binding, termed by the author as the "mapped 500" (Abruzzi et al., 2011). Among the 113 CLK targets I identified, 33 genes overlapped with those mapped 500 (Fig IV.5A), and the known CLK targets *vri*, *per*, *tim*, *cwo* and *pdp1*, were among the top 15 on our CLK target list. Remarkably, about 50% of the CLK targets overlap with CWO targets, and about 10% of the CWO targets overlap with CLK targets (Fig IV.5B), suggesting a potential functional connection between binding of these two transcription factors. Among the 55 overlapping targets are the known clock genes *per*, *tim*, *vri*, *Pdp1* and *cwo*, consistent with the result in Chapter III that CWO binds to these CLK targets and competes with CLK in E-box binding. These data, taken together, suggest that the competition between CLK and CWO binding reported in Chapter III could be a prominent pattern for global transcriptional regulation.

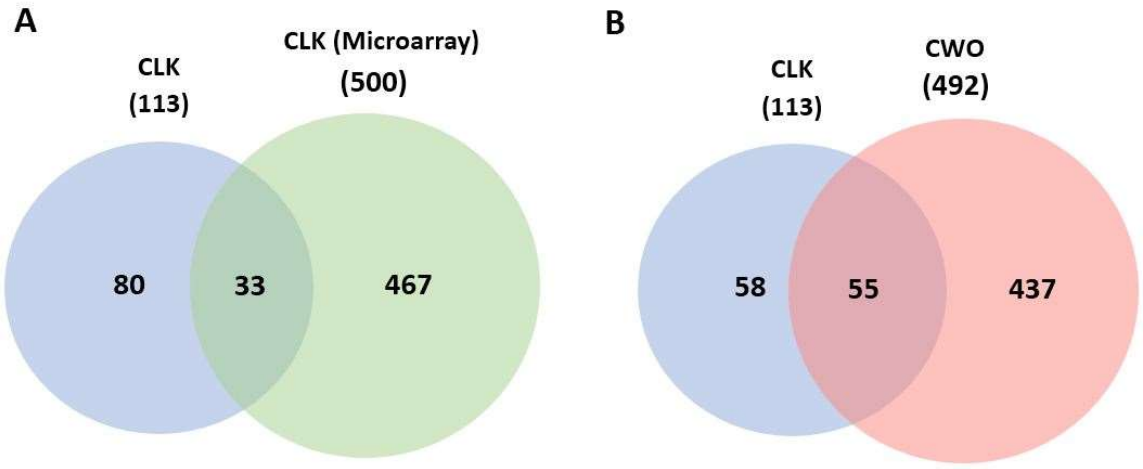


Fig IV.5. ChIP-seq analysis of CLK binding sites and comparison to CWO ChIP-seq and previous microarray data.

(A) Venn diagrams of comparison between CLK ChIP-seq targets (blue) and CLK microarray targets (green). **(B)** Venn diagrams of comparison between CLK ChIP-seq targets (blue) and CWO ChIP-seq targets (red). The numbers are determined as described in Fig IV.2A.

IV.B.3. Differential gene expression in wild-type vs. *cwo*^{5073} mutant flies

I then analyzed the transcriptome of wild-type (w^{1118}) and *cwo*^{5073} mutant flies at 6 time points from ZT2 to ZT22 in a LD cycle by RNA-seq and detected 15,175 RNA transcripts. From this total RNA transcripts, I considered only 11,778 RNAs whose level of expression (number of reads per kilobase per million mapped reads, RPKM) was >1 to avoid false differential expression caused by oscillations around a low basal expression. An arbitrary threshold was set to define CWO-regulated genes that resulted in a >30% decrease (*cwo*^{5073}/WT ratio <0.7) or a >50% increase (*cwo*^{5073}/WT ratio >1.5)

in expression in the absence of CWO. I found 588 overexpressed genes in which the *cwo*^{5073}/WT RPKM expression ratio was >1.5 and 582 downregulated genes in which the ratio was <0.7 (Fig IV.6A, B). GO analyses were performed to determine biological process and molecular functional categories of these 1170 differentially expressed genes. For biological process analysis, genes are enriched in metabolic processes and developmental processes, consistent with results from CWO ChIP-seq GO analyses. Genes involved in cellular process are also enriched, particularly those controlling cell communication and the cell cycle (Fig IV.6C). For molecular functional analysis, genes are enriched for the binding and catalytic activity categories, which contained DNA binding proteins such as transcriptional factors and enzymes involved in multiple biological processes, respectively (Fig IV.6D). These data suggest that CWO may potentially control gene expression via transcriptional regulation and posttranscriptional modifications at genome-wide level.

To determine whether CWO directly regulates downstream gene expression, I compared RNA-seq data with CWO ChIP-seq data described earlier. Of the direct target genes bound by CWO, I found 35 in which the *cwo*^{5073}/WT RPKM expression ratio was >1.5, and only 6 genes in which the ratio was <0.7 (Fig IV.6 A, B), indicating that the most frequent direct effect of the absence of CWO was an upregulation of gene expression. This observation is consistent with the role of CWO as a transcriptional repressor (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008; Zhou et al., 2016). However, majority of the differentially expressed genes are not CWO binding targets, suggesting that CWO widely regulates gene expression indirectly.

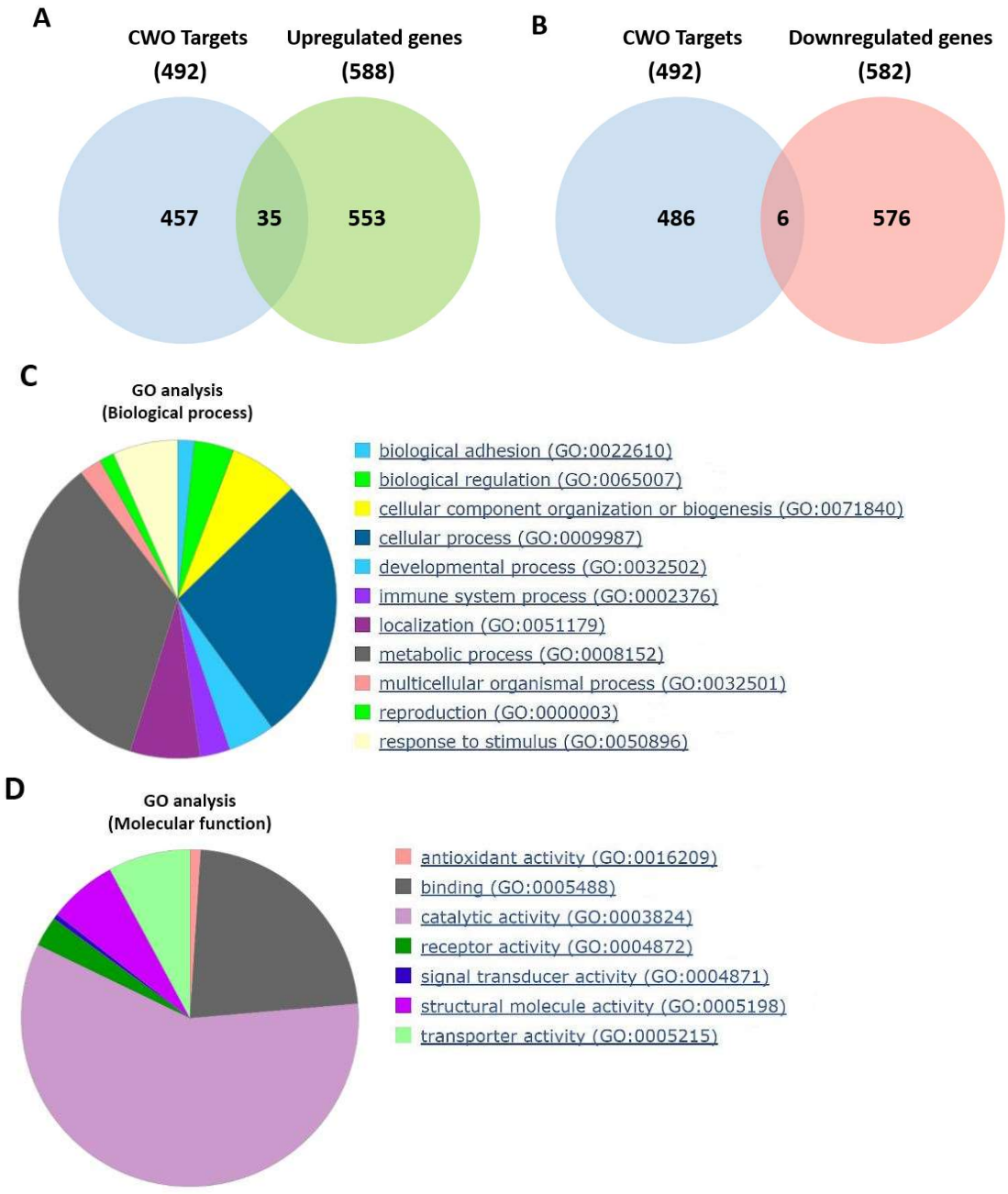


Fig IV.6. RNA-seq in WT vs. *cwo*^{5073} mutant identified differentially expressed mRNAs.

(A) Venn diagrams showing the overlap between CWO ChIP-seq targets (blue) and upregulated genes in *cwo*^{5073} mutant (green) **(B)** Venn diagrams showing the overlap

between CWO ChIP-seq targets (blue) and downregulated genes in *cwo*^{5073} mutant (red). The numbers are determined as described in Fig IV.2A. **(C)** Pie chart displays GO (biological process) analysis of differentially expressed genes in *cwo*^{5073} mutant. **(D)** Pie chart displays GO (molecular function) analysis of differentially expressed genes in *cwo*^{5073} mutant.

Upregulated	**Kinases**	***trbl, Hex-C***
	Phosphatases	***CG3264, CG3290, CG9449, CG11425, I-3, CG10592, CG8147, CG32568***
Downregulated	**Kinases**	***cdc2, CG5144, Rootletin, fj, CycB, CycA, Cks30A, p38c***
	Phosphatases	***CG17746***

Table IV.3. Kinases and phosphatases that are upregulated or downregulated in *cwo*^{5073} mutant.

IV.B.4. CWO potentially regulates CLK phosphorylation to promote transcription

A subset of differentially expressed genes fall into kinase and phosphatase GO categories (Table IV.3), suggesting that CWO is potentially involved in posttranscriptional regulation within the circadian feedback loops. Phosphorylation of PER regulates its nuclear entry and degradation, thus is critical for the timing of the circadian feedback (Chiu et al., 2008, 2011; Yu et al., 2011). CLK phosphorylation

coincides with transcriptional repression, suggesting the potential role of this CLK modification in transcriptional regulation (Kim and Edery, 2006; Yu et al., 2006, 2011). The phosphorylation states of CLK and PER can be determined by western blot analysis, as there are multiple electrophoretic mobility isoforms of PER and CLK arising from differential phosphorylation (Kim and Edery, 2006; Kim et al., 2007; Lee et al., 1998; Yu et al., 2006). Therefore, to determine whether CWO plays a role in the phosphorylation of PER or CLK, western analyses were carried out using head extracts from wild-type and *cwo*^{5073} flies collected every 4 hours in a LD cycle for PER, and at ZT2 and ZT14 for CLK. No detectable differences in PER electrophoretic mobility were observed at any time of day, suggesting that CWO does not affect PER phosphorylation (Fig IV.7A). CLK was hyperphosphorylated at ZT2 in both *cwo*^{5703} and wild-type flies, but at ZT14, the time when CLK is hypophosphorylated in wild-type flies, CLK remained predominantly hyperphosphorylated in the *cwo*^{5703} mutant (Fig IV.7B), indicating that the absence of CWO leads to CLK hyperphosphorylation. These results are consistent with previous observations that mRNA levels of the core clock genes were reduced in *cwo* RNAi knockdown and mutant flies (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008), probably because hyperphosphorylated CLK protein detected during early morning is associated with transcriptional repression (Kim and Edery, 2006; Yu et al., 2006). Based on my RNA-seq analysis, the kinase(s) or phosphatase(s) that may responsible for the change in CLK phosphorylation in *cwo*^{5073} flies are *trbl*, *Hex-C* and *CG17746,* because only upregulated kinases or downregulated phosphatases will cause hyperphosphorylation of CLK (Table

IV.3, Fig IV.8A-C). The levels of *trbl* and *Hex-C* mRNAs are increased at all time points during LD, while the levels of *CG17746* mRNA are decreased for most of the time points throughout the LD cycle (Fig IV.8A-C). Notably, ChIP-seq data shows no obvious binding of CWO or CLK to these kinase or phosphatase genes, suggesting an indirect effect of CWO on the mRNA levels of these candidates (Fig IV.8D-F). Among these candidate genes, *Hex-C* encodes a major form of hexokinase, an enzyme that phosphorylates hexose sugars, which makes it less likely to be a kinase that phosphorylates CLK. Further behavior and molecular analysis will provide more information, such as whether and how CLK phosphorylation is regulated by these kinase or phosphatase, and thus shed light on our understanding of how CWO controls the expression of clock genes at the posttranscriptional level.

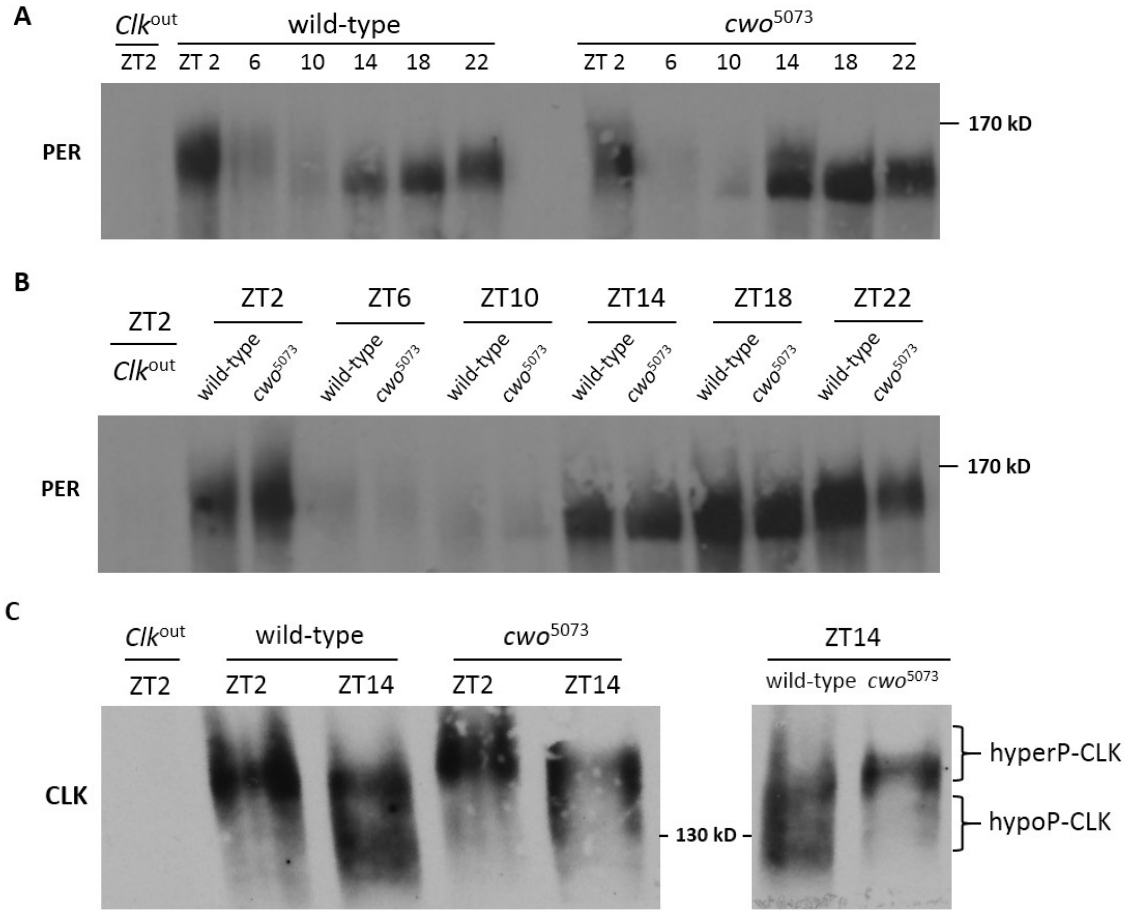


Fig IV.7. Phosphorylation of PER and CLK protein in wild-type and *cwo*^{5073} transgenic flies at different time points of the circadian cycles.

(A) Western blot of head extracts from wild-type and *cwo*^{5073} flies collected at the indicated times were probed with PER antiserum. Head extracts from *Clk*^{out} flies was used as negative control. **(B)** A repeat of the western blot in (A) with samples collected at the same times from wild-type and *cwo*^{5073} extracts loaded adjacently for head-to-head comparison. **(C)** Western blot of head extracts from wild-type and *cwo*^{5073} flies probed with CLK antiserum at ZT2 and ZT14 (left panel). Head extracts collected at ZT14 from each genotype were loaded adjacently for head-to-head comparison (right panel). Hyperphosphorylated CLK (hyperP-CLK) and hypophosphorylated CLK (hypoP-CLK) run as broad bands at 150 kDa and 120 kDa, respectively, indicated by the right braces on the right panel.

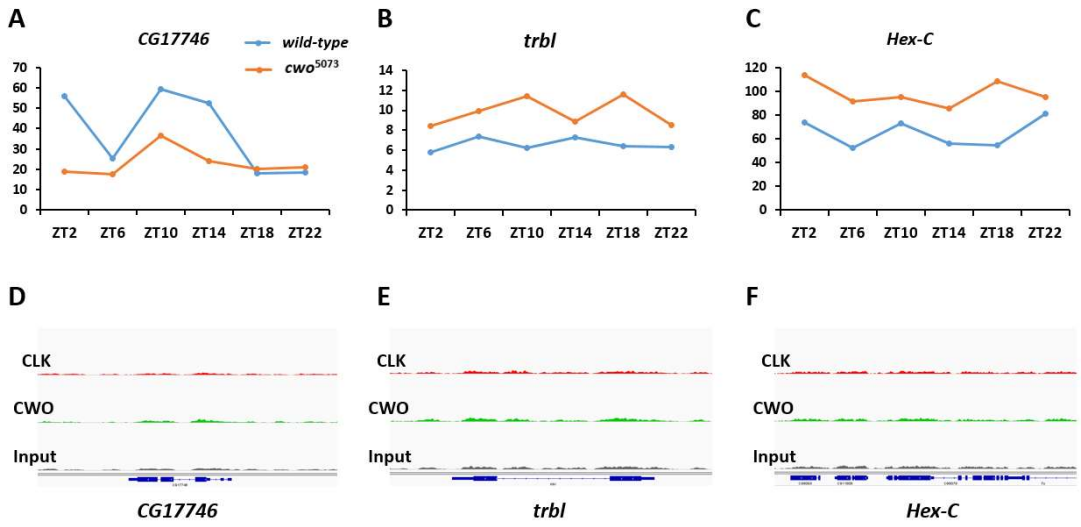


Fig IV.8. CWO indirectly regulates the mRNA levels of *CG17746*, *trbl* and *Hex-C*. **(A-C)** RNA-seq data showing the mRNA levels of *CG17746* **(A)**, *trbl* **(B)** and *Hex-C* **(C)** in wild-type (blue lines) and *cwo*^{5073} mutant (orange lines) strains. This RNA-seq experiments were only performed for one time so there is no error bar for the data. **(D-F)** ChIP-seq data showing that CLK (red) and CWO (green) do not bind *CG17746* **(D)**, *trbl* **(E)** and *Hex-C* **(F)** genomic regions. Gray color indicates input as a negative control for ChIP-seq.

IV.C. Summary

In this chapter I first demonstrate that CWO binds 492 genes by ChIP-seq analysis. A substantial overlap between CWO and CLK target genes is found, suggesting that CWO plays a potential role in regulating CLK-mediated transcription globally. This hypothesis is further supported by motif analysis showing that the canonical E-box sequences, the main binding sites of CLK, are highly enriched among CWO target regions. Further characterization of CWO temporal binding pattern and comparison of binding intensity

between CWO and CLK at genomic level will provide more information to uncover the general mechanism of CWO-mediated transcriptional repression in circadian clock.

In the second half of this chapter, I identify 588 upregulated genes and 582 downregulated genes in the *cwo*^{5073} mutant using RNA-seq, with only a few CWO direct target in common, suggesting that CWO mainly regulates gene expression in an indirect manner. Importantly, a subset of genes encoding kinases and phosphatases were found differentially expressed in *cwo*^{5073} mutant, which correlate with hyperphosphorylated CLK found in the mutants during the activation phase. However, it is not yet known whether these kinase or phosphatase are responsible for the change of CLK phosphorylation state. Further investigation of the molecular connection between CLK phosphorylation and the function of these kinases and phosphatases will help in understanding the role of CWO in posttranscriptional regulation of circadian feedback loops.

CHAPTER V

DISCUSSION AND SUMMARY

V.A. CWO competes with CLK in E-box binding to enhance PER mediated rhythms in transcriptional repression*

Rhythmic binding of CLK-CYC to E-boxes is essential for rhythmic transcription of the core circadian oscillator genes *per* and *tim* in *Drosophila*. CLK-CYC bind E-boxes upstream of *per* and *tim* in the late day and early night to activate transcription; and is released from these binding sites during late night (Menet et al., 2010; Taylor and Hardin, 2008; Yu et al., 2006). Previous work demonstrated that CLK constitutively binds *per* and *tim* E-boxes in per^{01} flies, indicating that PER is essential for rhythmic binding of CLK-CYC, and is key to removing CLK-CYC from E-boxes (Yu et al., 2006). In this study I report that CWO also contributes to removing CLK-CYC from the E-boxes. In cwo^{5703} mutant, CLK binding intensity is significantly increased at the trough of its binding cycle, suggesting that repression is incomplete in the absence of CWO (Fig III.4).

I find that CWO and CLK bind E-boxes upstream of *tim* in a reciprocal manner during a daily cycle, and that CLK shows significantly increased binding intensity at the trough of its binding cycle in *cwo* mutant flies, indicating that CWO acts to antagonize

---

*** 3. This part is partially reprinted with permission from Zhou J, Yu W, Hardin PE (2016) CLOCKWORK ORANGE Enhances PERIOD Mediated Rhythms in Transcriptional Repression by Antagonizing E-box Binding by CLOCK-CYCLE. PLoS Genet 12(11): e1006430. Copyright [2016] by PLOS.**

CLK-CYC binding. Given that both CWO and CLK are constitutively expressed (Fig III.1; Yu et al., 2006), I believe that the key driver for the transition between dynamic CLK-CYC and CWO binding is the accumulation of PER, which alters the relative affinity of E-box binding by CLK-CYC. CWO shows little or no *tim* E-box binding in per^{01} flies, in which CLK-CYC constantly bind E-boxes, but shows high levels of *tim* E-box binding in Clk^{out} flies that lack CLK expression and E-box occupancy. These results suggest that CWO E-box binding affinity is lower than the CLK-CYC heterodimer and higher than the CLK-CYC-PER complex. This intermediate binding affinity suggests a model for PER-dependent rhythms in CLK-CYC and CWO binding (Fig III.6). During late day and early night, CLK-CYC bind E-boxes to activate transcription because CLK-CYC has higher DNA binding affinity than CWO, which is present but cannot compete with CLK-CYC for binding. Later during the night PER starts to accumulate in the nucleus and interacts with CLK-CYC, which decreases CLK-CYC DNA interaction via reduced DNA binding affinity. Consequently, CWO out-competes CLK-CYC-PER by binding tandem E-box sites with comparatively higher affinity to release CLK-CYC-PER from the DNA. Once CLK-CYC-PER is removed, CWO occupancy on E-boxes prohibits CLK-CYC-PER from re-binding, thus maintaining transcriptional repression (Fig III.6). Co-expression of CWO and PER increases repression two to five fold more than either CWO or PER alone (Kadener et al., 2007), which is consistent with our model that CWO and PER work synergistically to repress CLK-CYC-mediated gene expression.

Unlike the constitutive CLK-CYC E-box binding in per^{01} flies (Yu et al., 2006),

CLK-CYC binding is rhythmic in *cwo*^{5703} flies, but with a dampened amplitude due to elevated CLK binding at the trough (Fig III.4A). This low amplitude rhythm in CLK binding may explain why a large proportion of *cwo*^{5703} flies show long period rhythms rather than losing rhythmicity entirely like *per*^{01} mutants (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). I speculate that the long period phenotype is caused in part by a prolonged repression process. Based on the current model for repression of CLK-CYC-mediated transcription, PER-TIM complexes first bind CLK-CYC, thereby removing CLK-CYC from the E-boxes and inhibiting *per* and *tim* transcription, then PER and TIM degradation enables CLK-CYC binding to start another cycle of transcription (Hardin, 2011). Both of these steps could be delayed in a *cwo* mutant. In the absence of CWO it takes longer to remove CLK-CYC from the DNA; PER alone can repress CLK-CYC binding to some degree, but CLK-CYC-PER complexes still weakly bind E-boxes if CWO is absent, thus reducing CLK-CYC repression compared to wild-type flies. The outcome of incomplete repression of CLK-CYC E-box binding would be an increase in the trough levels of *per* and *tim* mRNAs, which is exactly what was observed in *cwo* mutant and RNAi knockdown strains (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). Higher *per* and *tim* mRNA levels would in turn increase PER and TIM expression during the repression phase (Richier et al., 2008). Higher levels of PER and TIM would not repress CLK-CYC binding efficiently in the absence of CWO, but would take longer to be degraded, thereby delaying the next cycle of transcriptional activation.

In addition to the increased trough levels of core clock gene mRNAs in *cwo*

mutant and RNAi knockdown flies, the peak levels of these mRNAs are lower, particularly during DD (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). Decreasing *per* mRNA levels also lengthen circadian period (Baylies et al., 1987), thus making it difficult to determine the extent to which a lower mRNA peak or increased mRNA trough contributes to period lengthening in *cwo* mutant and RNAi knockdown flies. CLK binding at the peak of transcription is not significantly lower in cwo^{5073} than wild-type during LD (Fig III.4B), which argues that CWO enhances CLK-CYC transcriptional activity independent of CLK-CYC E-box binding. Additional experiments will be needed to decipher the mechanism underlying this CWO dependent increase in CLK-CYC transcription.

PER dependent repression of CLK-CYC transcription is thought to occur in two stages. First, PER is recruited to circadian promoters by interacting with CLK to form PER-CLK-CYC complexes "on-DNA", which inhibit CLK-CYC dependent transcription via an unknown mechanism. Subsequently, a decrease in the DNA binding affinity of PER-CLK-CYC complexes results in their release from DNA to initiate "off-DNA" phase of repression (Menet et al., 2010). According to our model, CWO is critical for the transition to, and maintenance of, off-DNA repression. When PER-CLK-CYC complexes with low DNA affinity are formed, CWO promotes off-DNA repression by competing with CLK-CYC-PER complex for E-box binding. CWO occupancy on E-boxes then prevents PER-CLK-CYC from re-binding, thereby maintaining off-DNA repression.

In mammals, a similar pattern of antagonistic binding on E-boxes between transcription factors was recently reported: USF1 and a mutant form of CLOCK, CLOCK$^{\Delta 19}$, bind to the same tandem E-boxes in a reciprocal manner. Wild-type CLOCK-BMAL1 complex binds E-boxes with much higher affinity than USF1, but CLOCK$^{\Delta 19}$-BMAL1 binds E-boxes with a similar affinity to USF1, thus allowing USF1 to bind E-boxes (Shimomura et al., 2013). Although this competitive binding is not thought to impact feedback loop function under normal circumstances, it demonstrates that other transcription factors can out-compete CLOCK-BMAL1 for E-box binding if the DNA binding affinity of CLOCK-BMAL1 is reduced. In this case CLOCK-BMAL1 binding is compromised by the *Clock*$^{\Delta 19}$ mutation, but other mechanisms such as interactions with repressors and protein modifications could also reduce the binding affinity of CLOCK-BMAL1 or its orthologs.

As in *Drosophila*, rhythmic binding of CLOCK-BMAL1 to E-boxes drives circadian transcription in mammals (reviewed in Gustafson and Partch, 2015). Recent ChIP-seq analyses in mouse liver revealed time-dependent binding of CLOCK, BMAL1 and key negative feedback components including PER1, PER2, CRY1 and CRY2 (Hatanaka et al., 2010; Koike et al., 2012; Menet et al., 2012; Rey et al., 2011). The mechanism underlying the dynamic DNA occupancy of these transcription factors is not known, but previous work shows that the PER2-CLOCK interaction is required to initiate repression of CLOCK-BMAL1 dependent transcription (Chen et al., 2009), which suggests that CLOCK-BMAL1 may be removed from E-boxes by the same mechanism as CLK-CYC in *Drosophila*. A recent genome-wide nucleosome analysis in

mouse liver revealed that rhythmic E-box binding by CLOCK-BMAL1 removes nucleosomes (Menet et al., 2014). However, despite rhythmic CLOCK-BMAL1 binding, nucleosome occupancy on E-boxes is always well below surrounding sequences, even in $Bmal1^{-/-}$ mutant livers (Menet et al., 2014). This result indicates that chromatin at CLOCK-BMAL1 target sites is not closed even when there is no CLOCK-BMAL1 binding, suggesting that other transcription factors may occupy these E-boxes when CLOCK-BMAL1 is absent. These results, taken together, suggest that rhythms in activator binding may be controlled by a common mechanism in *Drosophila* and mammals.

The mammalian orthologs of CWO, called DEC1 and DEC2 (and also SHARP2 and SHARP1, respectively), suppress CLOCK-BMAL1-induced activation (Bode et al., 2011; HAMAGUCHI et al., 2004; Honma et al., 2002b; LI et al., 2004b; Nakashima et al., 2008; Rossner et al., 2008). Gel mobility shift and ChIP assays *in vitro* revealed that both DEC1 and DEC2 bind to E-box motifs targeted by CLK-BMAL1 (Bode et al., 2011; Hamaguchi et al., 2004; Honma et al., 2002b; Li et al., 2004b; Nakashima et al., 2008), and the DNA-binding domain is required for DEC1 to regulate CLK-BMAL1-induced transactivation (Li et al., 2004b). In addition, DEC1/2 show synergistic activity to PER1 in the regulation of clock gene mRNA levels in the SCN, as exemplified by significant changes in the period of circadian activity rhythms when null mutants for *Dec1*, *Dec2* or both *Dec1* and *Dec2* are combined with that for *Per1* (Bode et al., 2011). These results raise the possibility that DEC1 and DEC2 may be a functional counterpart

of CWO in competing with CLOCK-BMAL1 for E-box binding to repress CLOCK-BMAL1-mediated transcription.

V.B. Genome-wide analysis of *cwo* function in transcriptional regulation

To investigate how transcription factor CWO regulates gene expression at genome-wide level, I carried out ChIP-seq and RNA-seq assays to search for direct and indirect CWO targets. By ChIP-seq I identified ~500 CWO direct target genes that have canonical E-box motifs highly enriched within the binding regions. Further RNA-seq analysis between wild-type and cwo^{5073} mutant flies identified about 1170 differentially expressed genes. Half of the genes are upregulated in cwo^{5073} flies and the other half are downregulated. These genome-wide analyses provide a rich resource to unravel the role of CWO in transcriptional and posttranscriptional gene regulation, therefore further improved our understanding of how CWO impacts circadian feedback loop function. For example, I find that CWO indirectly controls the expression of kinases and phosphatases that potentially regulate the phosphorylation state of CLK during transcriptional activation, thus acting to promote transcription of clock genes.

My ChIP-seq analysis identified 325 and 437 CWO direct target genes at ZT2 and ZT14, respectively. Motif analysis indicates high enrichment of the canonical CACGTG E-box motif among the mapped peaks, coinciding with earlier studies that CWO is a bHLH transcription factor that specifically binds E-boxes in cell culture and *in vivo* (Lim et al., 2007; Matsumoto et al., 2007; Zhou et al., 2016). Importantly, the core clock genes, *per*, *tim*, *vri*, *Pdp1* and *cwo* itself, which were bound by CWO in ChIP-

qPCR assays in Chapter III (Fig III.2), rank high on my ChIP-seq target list. GO analysis of CWO DNA-binding sites shows that basic metabolic and development-related functions were enriched among CWO targets. As expected, a group of genes controlling behavior are also enriched, including genes related to rhythmic processes and sleep, which coincides with the role of CWO in circadian regulation. Functionally, DNA binding proteins such as transcription factors are enriched among CWO targets, raising the possibility that CWO indirectly regulates global gene expression by directly activating or repressing the expression of these transcription factors. Interestingly, I find a substantial overlap between CLK and CWO binding sites by combining the data from our CLK and CWO ChIP-seq analyses. About 50% of CLK targets are also bound by CWO, suggesting that the competition for E-box binding between these two transcription factors could be a prominent pattern for circadian transcriptional repression. A more detailed characterization of CWO binding at multiple time points throughout a circadian cycle, and quantitative comparison of temporal binding intensity will be necessary to provide more information to support this hypothesis.

A previous ChIP-on-chip experiment had shown that CWO binds 1103 target genes in S2 cell culture, about 2 to 3 folds more than what I found *in vivo*. A likely reason for this difference is that CWO was overexpressed in the ChIP-on-chip experiment, which potentially leads to false-positive interactions since a much higher abundance of CWO could bind to DNA regions that are not bound by endogenous CWO protein. High levels of CWO could also out-compete other transcription factors for E-box binding, and therefore identify some sites artificially. Moreover, because of the

differences in the cellular and genomic environment between cultured cells and flies, cell culture studies may lead to results that do not correspond to the circumstances occurring in a living organism. One example was that the core clock gene *per* and *tim* were shown not bound by CWO in this ChIP-on-chip assay carried out in cell culture.

In Chapter III I show that CWO rhythmically binds E-boxes of the core clock genes, with a significantly higher binding intensity at ZT2. By ChIP-seq analysis I find CWO binds to these targets in a slightly different pattern which is somewhat different than that from ChIP-qPCR experiments (Fig IV.3). One possible reason for this inconsistency is that for ChIP-seq the fusion protein CWO-HA, instead of endogenous CWO, was immunoprecipitated. Although the expression pattern of CWO-HA is similar to the endogenous CWO, the MYC and HA tag sequence at the C-terminus of the protein could potentially interfere with protein function and DNA binding affinity. Indeed, the transgenic *cwo-HA; cwo*^{5073} flies does not fully rescue the long period phenotype observed in *cwo*^{5073} mutant, showing a behavior rhythm approximately 1 hour longer than wild-type flies (Table.IV.1). Also, possible PCR duplication during DNA library preparation could result in bias towards some targets at specific time points. Therefore, statistical analysis from more biological replicates of ChIP-seq experiments with increased sequencing depth would be helpful for quantitative binding analysis, which, as discussed earlier, will be important for the characterization of temporal binding pattern of CWO. With this additional information, we could then determine whether the competition binding model between CLK and CWO is a prominent mechanism for transcription repression.

As CWO was first reported as a transcriptional repressor (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007), I expected to see that more genes were upregulated in *cwo*^{5073} mutant by RNA-seq. However, the number of genes that were upregulated is approximately the same as the downregulated genes (Fig IV.6A, B). Combining RNA-seq with ChIP-seq data, I find that 35 upregulated genes are also CWO direct targets, but only 6 downregulated genes are directly bound by CWO, indicating that CWO mainly binds DNA to repress, but not activate, target genes expression. The overlap between CWO ChIP-seq and RNA-seq data is quite low, suggesting that most of the differential gene expression are caused by secondary or indirect effects. This idea is supported by the observation that a large proportion of CWO direct targets are transcription factors. Therefore, the differentially expressed genes in *cwo* mutant that are not bound CWO, especially the downregulated genes, are likely regulated by CWO indirectly. For example, CWO could bind to and transcriptionally repress a repressor, thus promoting activation of downstream genes. These data, taken together, suggest that CWO is a transcription factor that could directly or indirectly regulate transcriptional activation or repression globally.

Throughout the circadian cycle, CLK protein levels remain constant, and CLK-mediated transcription of target genes is dependent upon CLK protein modifications such as phosphorylation (Houl et al., 2006; Yu et al., 2006). CLK is hyperphosphorylated around dawn during transcription repression and hypophosphorylated around dusk during transcription activation (Yu et al., 2006). Although it is not clear whether CLK phosphorylation is the cause for repression, the

change of CLK phosphorylation state coincides with transcriptional activation and repression (Kim and Edery, 2006; Yu et al., 2009, 2011). Interestingly, I find CLK is hyperphosphorylated in *cwo*^{5073} flies during transcription activation at ZT14, which correlates with its reduced transcriptional activity since the mRNA levels of CLK target genes were much lower in *cwo*^{5073} flies (Kadener et al., 2007; Lim et al., 2007; Matsumoto et al., 2007; Richier et al., 2008). Using RNA-seq I identified genes encoding kinases and phosphatases that could potentially regulate CLK phosphorylation state were differentially expressed in *cwo*^{5073} flies. One of these kinase is *trbl*, which encodes the founding member of the Trib family of kinase-like proteins that regulate cell migration, proliferation, growth and homeostasis (reviewed in Dobens and Bouyain, 2012; Hegedus et al., 2007). However, it is not yet known whether this kinase is involved in circadian regulation, which could be further determined by behavior analysis of *trbl* RNAi knockdown of mutant strains. The mRNA level of *trbl* is increased in *cwo*^{5073} flies, coincident with hyperphosphorylation of CLK protein, making it a potential candidate CLK kinase. The other candidate is CG17746, a member of the protein phosphatase 2C family that has cation binding domains and dephosphorylates proteins at serine and threonine residues. The mRNA level of *CG17746* is reduced in *cwo*^{5073} flies, again coincident with hyperphosphorylation of CLK protein. When *CG17746* was specifically knocked down by RNAi in clock neurons, the transgenic flies have long behavior rhythm (Agrawal and Hardin, 2016), suggesting that *CG17746* potentially plays a role in the posttranscriptional regulation of circadian clock. Further validation of CWO's effect on the transcriptional and protein level of these candidate

genes, as well as characterization of the molecular connection between CLK phosphorylation and the function of these kinases and phosphatases will shed light on the understanding of CWO's role in posttranscriptional regulation of circadian feedback loops.

V.C. Conclusions

A major objective of my research is to determine how transcription factor CWO regulates the rhythmic transcription within the autoregulatory feedback loop that keeps circadian time in *Drosophila melanogaster*. The transcription repression process is important for the timing and stability of the circadian feedback loops, however the molecular mechanism largely remains unknown. In Chapter III, I demonstrate that CWO antagonizes CLK-CYC E-box binding, thus enhancing the removal of CLK-CYC from E-boxes to maintain transcriptional repression. This process requires PER, which suggests that PER-TIM and CWO cooperate to maintain a transcriptionally repressed state by removing CLK-CYC from E-boxes. These results demonstrate that PER-TIM require CWO to effectively repress circadian transcription, and given that circadian transcriptional regulators are well conserved, this mechanism may function to repress transcription in other animals including humans.

In Chapter IV, I carried out ChIP-seq and RNA-seq analyses and demonstrated that CWO widely binds to the genomic DNA and regulates gene expression in both direct and indirect manners. GO analyses suggest that CWO mainly involves in the regulation of metabolic and developmental process, and not surprisingly, behavior

control including circadian rhythm. About half of the CLK direct targets are also found as CWO targets by ChIP-seq, raising the possibility that CWO widely regulates CLK-mediated transcription, as shown in Chapter III, in a broader way. A large portion of CWO direct targets are transcription factors, and majority of the genes that differentially expressed in *cwo* mutant are not CWO direct targets, suggesting that CWO regulates gene expression mainly through indirect effect. This transcriptional network thus allows CWO to control the circadian feedback loops and potentially many other biological processes through various pathways at both transcriptional and posttranscriptional levels.

In the last part of Chapter IV, I demonstrate that CLK is hyperphosphorylated in cwo^{5073} mutant during the activation phase. Correlates with this result, a subset of genes encoding kinases and phosphatases were found differentially expressed in cwo^{5073} mutant, suggesting that CWO potentially controls circadian transcription at posttranscriptional level. Future works, such as behavior analyses of deficient or mutant flies to test whether those kinases and phosphatases are functionally clock related, validation of CWO's effect on the mRNA and protein levels of these candidate genes, and the characterization of the molecular mechanism of how these kinases and phosphatases regulate CLK phosphorylation will lead to a profound understanding of CWO's role in posttranscriptional regulation of circadian feedback loops.

REFERENCES

Abruzzi, K.C., Rodriguez, J., Menet, J.S., Desrochers, J., Zadina, A., Luo, W., Tkachev, S., and Rosbash, M. (2011). Drosophila CLOCK target gene characterization: implications for circadian tissue-specific gene expression. Genes Dev. *25*, 2374–2386.

Agrawal, P., and Hardin, P.E. (2016). An RNAi Screen To Identify Protein Phosphatases That Function Within the Drosophila Circadian Clock. G3 Genes, Genomes, Genet. *6*.

Allada, R., and Chung, B.Y. (2010). Circadian organization of behavior and physiology in Drosophila. Annu. Rev. Physiol. *72*, 605–624.

Allada, R., White, N.E., So, W.V., Hall, J.C., and Rosbash, M. (1998). A Mutant Drosophila Homolog of Mammalian Clock Disrupts Circadian Rhythms and Transcription of period and timeless. Cell *93*, 791–804.

Bae, K., Lee, C., Sidote, D., Chuang, K.Y., and Edery, I. (1998). Circadian regulation of a Drosophila homolog of the mammalian Clock gene: PER and TIM function as positive regulators. Mol. Cell. Biol. *18*, 6142–6151.

Baylies, M.K., Bargiello, T.A., Jackson, F.R., and Young, M.W. (1987). Changes in abundance or structure of the per gene product can alter periodicity of the Drosophila clock. Nature *326*, 390–392.

Bell-Pedersen, D., Cassone, V.M., Earnest, D.J., Golden, S.S., Hardin, P.E., Thomas, T.L., and Zoran, M.J. (2005). Circadian rhythms from multiple oscillators: lessons from diverse organisms. Nat. Rev. Genet. vol: 6 (7) pp: 544-556

Bischof, J., and Basler, K. (2008). Recombinases and their use in gene activation, gene inactivation, and transgenesis. Methods Mol. Biol. *420*, 175–195.

Blau, J., and Young, M.W. (1999). Cycling vrille Expression Is Required for a Functional Drosophila Clock. Cell *99*, 661–671.

Bode, B., Shahmoradi, A., Rossner, M.J., and Oster, H. (2011). Genetic Interaction of *Per1* and *Dec1/2* in the Regulation of Circadian Locomotor Activity. J. Biol. Rhythms *26*, 530–540.

Boy, A.L., Zhai, Z., Habring-Müller, A., Kussler-Schneider, Y., Kaspar, P., and Lohmann, I. (2010). Vectors for efficient and high-throughput construction of fluorescent drosophila reporters using the PhiC31 site-specific integration system. Genesis *48*, 452–456.

Busza, A., Emery-Le, M., Rosbash, M., and Emery, P. (2004). Roles of the Two Drosophila CRYPTOCHROME Structural Domains in Circadian Photoreception. Science (80-.). *304*.

Ceriani, M.F., Darlington, T.K., Staknis, D., Más, P., Petti, A.A., Weitz, C.J., and Kay, S.A. (1999). Light-Dependent Sequestration of TIMELESS by CRYPTOCHROME. Science (80-.). *285*.

Chaix, A., Zarrinpar, A., Miu, P., and Panda, S. (2014). Time-Restricted Feeding Is a Preventative and Therapeutic Intervention against Diverse Nutritional Challenges. Cell Metab. *20*, 991–1005.

Chang, D.C., and Reppert, S.M. (2003). A Novel C-Terminal Domain of Drosophila PERIOD Inhibits dCLOCK:CYCLE-Mediated Transcription. Curr. Biol.*13, 758-762*

Chen, R., Schirmer, A., Lee, Y., Lee, H., Kumar, V., Yoo, S.-H., Takahashi, J.S., and Lee, C. (2009). Rhythmic PER Abundance Defines a Critical Nodal Point for Negative Feedback within the Circadian Clock Mechanism. Mol. Cell *36*, 417–430.

Chen, W., Liu, Z., Li, T., Zhang, R., Xue, Y., Zhong, Y., Bai, W., Zhou, D., and Zhao, Z. (2014). Regulation of Drosophila circadian rhythms by miRNA let-7 is mediated by a regulatory cycle. Nat. Commun. *5*, 5549.

Chiu, J.C., Vanselow, J.T., Kramer, A., and Edery, I. (2008). The phospho-occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. Genes Dev. *22*, 1758–1772.

Chiu, J.C., Ko, H.W., and Edery, I. (2011). NEMO/NLK Phosphorylates PERIOD to Initiate a Time-Delay Phosphorylation Circuit that Sets Circadian Clock Speed. Cell *145*, 357–370.

Chung, S., Lee, E.J., Yun, S., Choe, H.K., Park, S.-B., Son, H.J., Kim, K.-S., Dluzen, D.E., Lee, I., Hwang, O., et al. (2014). Impact of Circadian Nuclear Receptor REV-ERBα on Midbrain Dopamine Production and Mood Regulation. Cell *157*, 858–868.

Claridge-Chang, A., Wijnen, H., Naef, F., Boothroyd, C., Rajewsky, N., and Young, M.W. (2001). Circadian Regulation of Gene Expression Systems in the Drosophila Head. Neuron *32*, 657–671.

Curtin, K.D., Huang, Z.J., and Rosbash, M. (1995). Temporally Regulated Nuclear Entry of the Drosophila period Protein Contributes to the Circadian Clock. Neuron *14*, 365–372.

Cyran, S.A., Buchsbaum, A.M., Reddy, K.L., Lin, M.-C., Glossop, N.R.J., Hardin, P.E., Young, M.W., Storti, R. V., and Blau, J. (2003). vrille, Pdp1, and dClock Form a Second Feedback Loop in the Drosophila Circadian Clock. Cell *112*, 329–341.

Darlington, T.K., Wager-Smith, K., Ceriani, M.F., Staknis, D., Gekakis, N., Steeves, T.D.L., Weitz, C.J., Takahashi, J.S., and Kay, S.A. (1998). Closing the Circadian Loop: CLOCK-Induced Transcription of Its Own Inhibitors per and tim. Science (80-.). *280*.

Darlington, T.K., Lyons, L.C., Hardin, P.E., and Kay, S.A. (2000). The period E-box Is Sufficient to Drive Circadian Oscillation of Transcription In Vivo. J. Biol. Rhythms *15*, 462–470.

Davis, R.L., and Turner, D.L. (2001). Vertebrate hairy and Enhancer of split related proteins: transcriptional repressors regulating cellular differentiation and embryonic patterning. Oncogene *20*, 8342–8357.

Dawson, S.R., Turner, D.L., Weintraub, H., and Parkhurst, S.M. (1995). Specificity for the hairy/enhancer of split basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression. Mol. Cell. Biol. *15*, 6923–6931.

Dissel, S., Codd, V., Fedic, R., Garner, K.J., Costa, R., Kyriacou, C.P., and Rosato, E. (2004). A constitutively active cryptochrome in Drosophila melanogaster. Nat. Neurosci. *7*, 834–840.

Dobens, L.L., and Bouyain, S. (2012). Developmental roles of tribbles protein family members. Dev. Dyn. *241*, 1239–1248.

Dunlap, J.C. (1999). Molecular Bases for Circadian Clocks. Cell *96*, 271–290.

Fisher, A.L., Ohsako, S., and Caudy, M. (1996). The WRPW Motif of the Hairy-Related Basic Helix-Loop-Helix Repressor Proteins Acts as a 4-Amino-Acid Transcription Repression and Protein-Protein Interaction Domain. Mol. Cell. Biol. *16*, 2670–2677.

Gekakis, N., Staknis, D., Nguyen, H.B., Davis, F.C., Wilsbacher, L.D., King, D.P., Takahashi, J.S., and Weitz, C.J. (1998). Role of the CLOCK Protein in the Mammalian Circadian Mechanism. Science (80-.). *280*.

Glossop, N.R.J., Lyons, L.C., and Hardin, P.E. (1999). Interlocked Feedback Loops Within the Drosophila Circadian Oscillator. Science (80-.). *286*.

Glossop, N.R.J., Houl, J.H., Zheng, H., Ng, F.S., Dudek, S.M., and Hardin, P.E. (2003). VRILLE Feeds Back to Control Circadian Transcription of Clock in the Drosophila Circadian Oscillator. Neuron *37*, 249–261.

Groth, A.C., Fish, M., Nusse, R., and Calos, M.P. (2004). Construction of transgenic Drosophila by using the site-specific integrase from phage phiC31. Genetics *166*, 1775–1782.

Gustafson, C.L., and Partch, C.L. (2015). Emerging Models for the Molecular Basis of Mammalian Circadian Timing. Biochemistry *54*, 134–149.

HAMAGUCHI, H., FUJIMOTO, K., KAWAMOTO, T., NOSHIRO, M., MAEMURA, K., TAKEDA, N., NAGAI, R., FURUKAWA, M., HONMA, S., HONMA, K., et al. (2004). Expression of the gene for Dec2, a basic helix–loop–helix transcription factor, is regulated by a molecular clock system. Biochem. J. *382*.

Hao, H., Allen, D.L., and Hardin, P.E. (1997). A Circadian Enhancer Mediates PER-Dependent mRNA Cycling in Drosophila melanogaster. Curr. Biol. *17*, 3687–3693.

Hao, H., Glossop, N.R.J., Lyons, L., Qiu, J., Morrish, B., Cheng, Y., Helfrich-Förster, C., and Hardin, P. (1999). The 69 bp Circadian Regulatory Sequence (CRS) Mediatesper-Like Developmental, Spatial, and Circadian Expression and Behavioral Rescue in Drosophila. J. Neurosci. *19*.

Hardin, P.E. (2004). Transcription Regulation within the Circadian Clock: The E-box and Beyond. J. Biol. Rhythms *19*, 348–360.

Hardin, P.E. (2005). The circadian timekeeping system of Drosophila. Curr. Biol. *15*, R714-22.

Hardin, P.E. (2011). Molecular genetic analysis of circadian timekeeping in Drosophila. Adv. Genet. *74*, 141–173.

Hardin, P.E., Hall, J.C., and Rosbash, M. (1990). Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature *343*, 536–540.

Hatanaka, F., Matsubara, C., Myung, J., Yoritaka, T., Kamimura, N., Tsutsumi, S., Kanai, A., Suzuki, Y., Sassone-Corsi, P., Aburatani, H., et al. (2010). Genome-wide profiling of the core clock protein BMAL1 targets reveals a strict relationship with metabolism. Mol. Cell. Biol. *30*, 5636–5648.

He, Y., Jones, C.R., Fujiki, N., Xu, Y., Guo, B., Holder, J.L., Rossner, M.J., Nishino, S., and Fu, Y.-H. (2009). The Transcriptional Repressor DEC2 Regulates Sleep Length in Mammals. Science (80-.). *325*.

Hegedus, Z., Czibula, A., and Kiss-Toth, E. (2007). Tribbles: A family of kinase-like proteins with potent signalling regulatory function. Cell. Signal. *19*, 238–250.

Helfrich-Förster, C. (2003). The neuroarchitecture of the circadian clock in the brain of *Drosophila melanogaster*. Microsc. Res. Tech. *62*, 94–102.

Hogenesch, J.B., Gu, Y.-Z., Moran, S.M., Shimomura, K., Radcliffe, L.A., Takahashi, J.S., and Bradfield, C.A. (2000). The Basic Helix-Loop-Helix-PAS Protein MOP9 Is a Brain-Specific Heterodimeric Partner of Circadian and Hypoxia Factors. J. Neurosci. *20*.

Honma, S., Kawamoto, T., Takagi, Y., Fujimoto, K., Sato, F., Noshiro, M., Kato, Y., and Honma, K. (2002a). Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature *419*, 841–844.

Honma, S., Kawamoto, T., Takagi, Y., Fujimoto, K., Sato, F., Noshiro, M., Kato, Y., and Honma, K. (2002b). Dec1 and Dec2 are regulators of the mammalian molecular clock. Nature *419*, 841–844.

Houl, J.H., Yu, W., Dudek, S.M., and Hardin, P.E. (2006). Drosophila CLOCK is constitutively expressed in circadian oscillator and non-oscillator cells. J. Biol. Rhythms *21*, 93–103.

Iyer, E.P.R., Iyer, S.C., Sullivan, L., Wang, D., Meduri, R., Graybeal, L.L., and Cox, D.N. (2013). Functional Genomic Analyses of Two Morphologically Distinct Classes of Drosophila Sensory Neurons: Post-Mitotic Roles of Transcription Factors in Dendritic Patterning. PLoS One *8*, e72434.

Kadener, S., Stoleru, D., McDonald, M., Nawathean, P., and Rosbash, M. (2007). Clockwork Orange is a transcriptional repressor and a new Drosophila circadian pacemaker component. Genes Dev. *21*, 1675–1686.

Kawamoto, T., Noshiro, M., Sato, F., Maemura, K., Takeda, N., Nagai, R., Iwata, T., Fujimoto, K., Furukawa, M., Miyazaki, K., et al. (2004). A novel autofeedback loop of Dec1 transcription involved in circadian rhythm regulation. Biochem. Biophys. Res. Commun. *313*, 117–124.

Kim, E.Y., and Edery, I. (2006). Balance between DBT/CKIepsilon kinase and protein phosphatase activities regulate phosphorylation and stability of Drosophila CLOCK protein. Proc. Natl. Acad. Sci. U. S. A. *103*, 6178–6183.

Kim, E.Y., Ko, H.W., Yu, W., Hardin, P.E., and Edery, I. (2007). A DOUBLETIME kinase binding domain on the Drosophila PERIOD protein is essential for its hyperphosphorylation, transcriptional repression, and circadian clock function. Mol. Cell. Biol. *27*, 5014–5028.

Kloss, B., Price, J.L., Saez, L., Blau, J., Rothenfluh, A., Wesley, C.S., and Young, M.W. (1998). The Drosophila Clock Gene double-time Encodes a Protein Closely Related to Human Casein Kinase Iε. Cell *94*, 97–107.

Kloss, B., Rothenfluh, A., Young, M.W., and Saez, L. (2001). Phosphorylation of PERIOD Is Influenced by Cycling Physical Associations of DOUBLE-TIME, PERIOD, and TIMELESS in the Drosophila Clock. Neuron *30*, 699–706.

Ko, H.W., Jiang, J., and Edery, I. (2002). Role for Slimb in the degradation of Drosophila Period protein phosphorylated by Doubletime. Nature *420*, 673–678.

Ko, H.W., Kim, E.Y., Chiu, J., Vanselow, J.T., Kramer, A., and Edery, I. (2010). A Hierarchical Phosphorylation Cascade That Regulates the Timing of PERIOD Nuclear

Entry Reveals Novel Roles for Proline-Directed Kinases and GSK-3β/SGG in Circadian Clocks. J. Neurosci. *30. 12664-75*

Koh, K., Zheng, X., and Sehgal, A. (2006). JETLAG Resets the Drosophila Circadian Clock by Promoting Light-Induced Degradation of TIMELESS. Science (80-.). *312*.

Koike, N., Yoo, S.-H., Huang, H.-C., Kumar, V., Lee, C., Kim, T.-K., and Takahashi, J.S. (2012). Transcriptional Architecture and Chromatin Landscape of the Core Circadian Clock in Mammals. Science (80-.). *338*.

Kume, K., Zylka, M.J., Sriram, S., Shearman, L.P., Weaver, D.R., Jin, X., Maywood, E.S., Hastings, M.H., Reppert, S.M., Takahashi, J.., et al. (1999). mCRY1 and mCRY2 Are Essential Components of the Negative Limb of the Circadian Clock Feedback Loop. Cell *98*, 193–205.

Lee, C., Parikh, V., Itsukaichi, T., Bae, K., and Edery, I. (1996). Resetting the Drosophila Clock by Photic Regulation of PER and a PER-TIM Complex. Science (80-.). *271*.

Lee, C., Bae, K., and Edery, I. (1998). The Drosophila CLOCK Protein Undergoes Daily Rhythms in Abundance, Phosphorylation, and Interactions with the PER–TIM Complex. Neuron *21*, 857–867.

Lee, C., Bae, K., and Edery, I. (1999). PER and TIM inhibit the DNA binding activity of a Drosophila CLOCK-CYC/dBMAL1 heterodimer without disrupting formation of the heterodimer: a basis for circadian transcription. Mol. Cell. Biol. *19*, 5316–5325.

LI, Y., SONG, X., MA, Y., LIU, J., YANG, D., and YAN, B. (2004). DNA binding, but not interaction with Bmal1, is responsible for DEC1-mediated transcription regulation of the circadian gene mPer1. Biochem. J. *382*.

Lim, C., Chung, B.Y., Pitman, J.L., McGill, J.J., Pradhan, S., Lee, J., Keegan, K.P., Choe, J., and Allada, R. (2007). clockwork orange Encodes a Transcriptional Repressor Important for Circadian-Clock Amplitude in Drosophila.

Lin, F.J., Song, W., Meyer-Bernstein, E., Naidoo, N., and Sehgal, A. (2001). Photic signaling by cryptochrome in the Drosophila circadian system. Mol. Cell. Biol. *21*, 7287–7294.

Lin, J.-M., Kilman, V.L., Keegan, K., Paddock, B., Emery-Le, M., Rosbash, M., and Allada, R. (2002a). A role for casein kinase 2α in the Drosophila circadian clock. Nature *420*, 816–820.

Lin, Y., Han, M., Shimada, B., Wang, L., Gibler, T.M., Amarakone, A., Awad, T.A., Stormo, G.D., Van Gelder, R.N., and Taghert, P.H. (2002b). Influence of the period-dependent circadian clock on diurnal, circadian, and aperiodic gene expression in Drosophila melanogaster. Proc. Natl. Acad. Sci. U. S. A. *99*, 9562–9567.

Littlewood, T.D., and Evan, G.I. (1995). Transcription factors 2: helix-loop-helix. Protein Profile *2*, 621–702.

Lowrey, P.L., and Takahashi, J.S. (2011). Chapter 6 – Genetics of Circadian Rhythms in Mammalian Model Organisms. In Advances in Genetics, pp. 175–230.

Lyons, L.C., Darlington, T.K., Hao, H., Houl, J., Kay, S.A., and Hardin, P.E. (2000). Specific Sequences Outside the E-box Are Required for Proper per Expression and Behavioral Rescue. J. Biol. Rhythms *15*, 472–482.

Mahesh, G., Jeong, E., Ng, F.S., Liu, Y., Gunawardhana, K., Houl, J.H., Yildirim, E., Amunugama, R., Jones, R., Allen, D.L., et al. (2014). Phosphorylation of the transcription activator CLOCK regulates progression through a ~ 24-h feedback loop to influence the circadian period in Drosophila. J. Biol. Chem. *289*, 19681–19693.

Martinek, S., Inonog, S., Manoukian, A.S., and Young, M.W. (2001). A Role for the Segment Polarity Gene shaggy/GSK-3 in the Drosophila Circadian Clock. Cell *105*, 769–779.

Massari, M.E., and Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. Mol. Cell. Biol. *20*, 429–440.

Matsumoto, A., Ukai-Tadenuma, M., Yamada, R.G., Houl, J., Uno, K.D., Kasukawa, T., Dauwalder, B., Itoh, T.Q., Takahashi, K., Ueda, R., et al. (2007). A functional genomics strategy reveals clockwork orange as a transcriptional regulator in the Drosophila circadian clock. Genes Dev. *21*, 1687–1700.

McDonald, M.J., and Rosbash, M. (2001). Microarray Analysis and Organization of Circadian Gene Expression in Drosophila. Cell *107*, 567–578.

McDonald, M.J., Rosbash, M., and Emery, P. (2001). Wild-type circadian rhythmicity is dependent on closely spaced E boxes in the Drosophila timeless promoter. Mol. Cell. Biol. *21*, 1207–1217.

Menet, J.S., Abruzzi, K.C., Desrochers, J., Rodriguez, J., and Rosbash, M. (2010). Dynamic PER repression mechanisms in the Drosophila circadian clock: from on-DNA to off-DNA. Genes Dev. *24*, 358–367.

Menet, J.S., Rodriguez, J., Abruzzi, K.C., and Rosbash, M. (2012). Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. Elife *1*, e00011.

Menet, J.S., Pescatore, S., and Rosbash, M. (2014). CLOCK:BMAL1 is a pioneer-like transcription factor. Genes Dev. *28*, 8–13.

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., and Thomas, P.D. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. *45*, D183–D189.

Moore, R.Y., and Eichler, V.B. (1972). Loss of a circadian adrenal corticosterone rhythm following suprachiasmatic lesions in the rat. *42, 201-206.*

Naidoo, N., Song, W., Hunter-Ensor, M., and Sehgal, A. (1999). A Role for the Proteasome in the Light Response of the Timeless Clock Protein. Science (80-.). *285*.

Nakashima, A., Kawamoto, T., Honda, K.K., Ueshima, T., Noshiro, M., Iwata, T., Fujimoto, K., Kubo, H., Honma, S., Yorioka, N., et al. (2008). DEC1 modulates the circadian phase of clock gene expression. Mol. Cell. Biol. *28*, 4080–4092.

Paquet, E.R., Rey, G., Naef, F., Zakany, J., and Duboule, D. (2008). Modeling an Evolutionary Conserved Circadian Cis-Element. PLoS Comput. Biol. *4*, e38.

Patke, A., Murphy, P.J., Onat, O.E., Krieger, A.C., Özçelik, T., Campbell, S.S., and Young, M.W. (2017). Mutation of the Human Circadian Clock Gene CRY1 in Familial Delayed Sleep Phase Disorder. Cell *169*, 203–215.e13.

Peschel, N., Chen, K.F., Szabo, G., and Stanewsky, R. (2009). Light-Dependent Interactions between the Drosophila Circadian Clock Factors Cryptochrome, Jetlag, and Timeless. *19 , 241-247.*

Pfeiffenberger, C., Lear, B.C., Keegan, K.P., and Allada, R. (2010). Processing sleep data created with the Drosophila Activity Monitoring (DAM) System. Cold Spring Harb. Protoc. *2010*, pdb.prot5520.

Preitner, N., Damiola, F., Luis-Lopez-Molina, Zakany, J., Duboule, D., Albrecht, U., and Schibler, U. (2002). The Orphan Nuclear Receptor REV-ERBα Controls Circadian Transcription within the Positive Limb of the Mammalian Circadian Oscillator. Cell *110*, 251–260.

Price, J.L., Blau, J., Rothenfluh, A., Abodeely, M., Kloss, B., and Young, M.W. (1998). double-time Is a Novel Drosophila Clock Gene that Regulates PERIOD Protein Accumulation. Cell *94*, 83–95.

Rey, G., Cesbron, F., Rougemont, J., Reinke, H., Brunner, M., and Naef, F. (2011). Genome-Wide and Phase-Specific DNA-Binding Rhythms of BMAL1 Control Circadian Output Functions in Mouse Liver. PLoS Biol. *9*, e1000595.

Richier, B., Michard-Vanhée, C., Lamouroux, A., Papin, C., and Rouyer, F. (2008). The Clockwork Orange Drosophila Protein Functions as Both an Activator and a Repressor of Clock Gene Expression. J. Biol. Rhythms *23*, 103–116.

Rossner, M.J., Oster, H., Wichert, S.P., Reinecke, L., Wehr, M.C., Reinecke, J., Eichele, G., Taneja, R., and Nave, K.-A. (2008). Disturbed Clockwork Resetting in Sharp-1 and Sharp-2 Single and Double Mutant Mice. PLoS One *3*, e2762.

Rutila, J.E., Suri, V., Le, M., So, W.V., Rosbash, M., and Hall, J.C. (1998). CYCLE Is a Second bHLH-PAS Clock Protein Essential for Circadian Rhythmicity and Transcription of Drosophila period and timeless. Cell *93*, 805–814.

Sato, T.K., Panda, S., Miraglia, L.J., Reyes, T.M., Rudic, R.D., McNamara, P., Naik, K.A., FitzGerald, G.A., Kay, S.A., and Hogenesch, J.B. (2004). A Functional Genomics Strategy Reveals Rora as a Component of the Mammalian Circadian Clock. Neuron *43*, 527–537.

Sehgal, A., Rothenfluh-Hilfiker, A., Hunter-Ensor, M., Chen, Y., Myers, M.P., and Young, M.W. (1995). Rhythmic Expression of timeless: A Basis for Promoting Circadian Cycles in period Gene Autoregulation. Science (80-.). *270*.

Shafer, O.T., Helfrich-Förster, C., Renn, S.C.P., and Taghert, P.H. (2006). Reevaluation of Drosophila melanogaster's neuronal circadian pacemakers reveals new neuronal classes. J. Comp. Neurol. *498*, 180–193.

Shimomura, K., Kumar, V., Koike, N., Kim, T.-K., Chong, J., Buhr, E.D., Whiteley, A.R., Low, S.S., Omura, C., Fenner, D., et al. (2013). *Usf1* , a suppressor of the circadian *Clock* mutant, reveals the nature of the DNA-binding of the CLOCK:BMAL1 complex in mice. Elife *2*, e00426.

Stanewsky, R., Kaneko, M., Emery, P., Beretta, B., Wager-Smith, K., Kay, S.A., Rosbash, M., Hall, J.C., Rosbash, M., and Hall, J.. (1998). The cryb mutation identifies cryptochrome as a circadian photoreceptor in Drosophila. Cell *95*, 681–692.

Stephan, F.K., and Zucker, I. (1972). Circadian rhythms in drinking behavior and locomotor activity of rats are eliminated by hypothalamic lesions. Proc. Natl. Acad. Sci. U. S. A. *69*, 1583–1586.

Taylor, P., and Hardin, P.E. (2008). Rhythmic E-box binding by CLK-CYC controls daily cycles in per and tim transcription and chromatin modifications. Mol. Cell. Biol. *28*, 4642–4652.

Ueda, H.R., Matsumoto, A., Kawamura, M., Iino, M., Tanimura, T., and Hashimoto, S. (2002). Genome-wide transcriptional orchestration of circadian rhythms in Drosophila. J. Biol. Chem. *277*, 14048–14052.

Venken, K.J.T., He, Y., Hoskins, R.A., and Bellen, H.J. (2006). P[acman]: A BAC Transgenic Platform for Targeted Insertion of Large DNA Fragments in D. melanogaster. Science (80-.). *314*.

Venken, K.J.T., Kasprowicz, J., Kuenen, S., Yan, J., Hassan, B.A., and Verstreken, P. (2008). Recombineering-mediated tagging of Drosophila genomic constructs for in vivo localization and acute protein inactivation. Nucleic Acids Res. *36*, e114–e114.

Wang, G.K., Ousley, A., Darlington, T.K., Chen, D., Chen, Y., Fu, W., Hickman, L.J., Kay, S.A., and Sehgal, A. (2001). Regulation of the cycling oftimeless (tim) RNA. J. Neurobiol. *47*, 161–175.

Xu, Y., Toh, K.L., Jones, C.R., Shin, J.-Y., Fu, Y.-H., Ptácek, L.J., Lee, C.C., Bradley, A., Fu, Y.-H., Sugano, S., et al. (2007). Modeling of a human circadian mutation yields insights into clock regulation by PER2. Cell *128*, 59–70.

Young, M.W., and Kay, S.A. (2001). Time zones: a comparative genetics of circadian clocks. Nat. Rev. Genet. *2*, 702–715.

Yu, W., Zheng, H., Houl, J.H., Dauwalder, B., and Hardin, P.E. (2006). PER-dependent rhythms in CLK phosphorylation and E-box binding regulate circadian transcription. Genes Dev. *20*, 723–733.

Yu, W., Zheng, H., Price, J.L., and Hardin, P.E. (2009). DOUBLETIME Plays a Noncatalytic Role To Mediate CLOCK Phosphorylation and Repress CLOCK-Dependent Transcription within the Drosophila Circadian Clock. Mol. Cell. Biol. *29*, 1452–1458.

Yu, W., Houl, J.H., and Hardin, P.E. (2011). NEMO Kinase Contributes to Core Period Determination by Slowing the Pace of the Drosophila Circadian Oscillator. Curr. Biol. *21, 756-761.*

Zeng, H., Qian, Z., Myers, M.P., and Rosbash, M. (1996). A light-entrainment mechanism for the Drosophila circadian clock. Nature *380*, 129–135.

Zhang, R., Lahens, N.F., Ballance, H.I., Hughes, M.E., and Hogenesch, J.B. (2014). A circadian gene expression atlas in mammals: Implications for biology and medicine. Proc. Natl. Acad. Sci. *111*, 16219–16224.

Zhou, J., Yu, W., and Hardin, P.E. (2015). Chipping away at the Drosophila clock. Methods in Enzymology *551. 323-347.*

Zhou, J., Yu, W., and Hardin, P.E. (2016). CLOCKWORK ORANGE Enhances PERIOD Mediated Rhythms in Transcriptional Repression by Antagonizing E-box Binding by CLOCK-CYCLE. PLoS Genet. *12*, 1–17.

APPENDIX

The solutions needed for ChIP in Chapter III are listed below.

XIP Homogenization Buffer (HB)*					
Component	Final Concentration	Stock Solution	3ml	5ml	10ml
Water			1.37ml	2.29ml	4.58ml
2x XIP-HB-HSEEIT	1x	2x	1.5ml	2.5ml	5ml
Formaldehyde (HCHO)	1%	37%	81.1µl	135.1µl	270.2µl
PMSF[a]	1mM	200mM @4°C	15µl	25µl	50µl
Na_3VO_4	1mM	200mM@-20°C	15µl	25µl	50µl
NaF	1mM	200mM@RT	15µl	25µl	50µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

XIP Homogenize Dilution Buffer (HDB)*				
Component	Final Concentration	Stock Solution	10ml	20ml
Water			4.84ml	9.67ml
2x XIP-HB-HSEEIT	1x	2x	5ml	10ml

PMSF[a]	0.5mM	200mM @4°C	25 µl	50µl
Aprotinin	10µg/ml	10µg/µl@4°C	10µl	20µl
Leupeptin	10µg/ml	10µg/µl@-20°C	10µl	20µl
Pepstatin A	2µg/ml	1µg/µl@-20°C	20µl	40µl
Na_3VO_4	1mM	200mM@-20°C	50µl	100µl
NaF	1mM	200mM@RT	50µl	100µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

2x XIP-HB-HSEEIT*					
Component	Final Concentration	Stock Solution	30ml	50ml	200ml
Water			21.525ml	35.88ml	143.5ml
HEPES•K (pH8.0)	100mM	1M @4°C	3ml	5ml	20ml
NaCl	280mM	5M @RT	1.68ml	2.8ml	11.2ml
EDTA (pH8.5)	2mM	0.5M @4°C	120µl	200µl	0.8ml
EGTA (pH8.0)	1mM	0.4M @4°C	75µl	125µl	0.5ml
Igpel CA-630	0.8%	10%	2.4ml	4ml	16ml
Triton X-100	0.4%	10%	1.2ml	2ml	8ml

* This solution can be stored temporary at @4°C, or long term @-20°C

XIP Nuclei Wash Buffer*						
Component	Final Concentration	Stock Solution	5 ml	10 ml	15 ml	20 ml
Water			4.4ml	8.835ml	13.215 ml	17.67ml
10x XIP-TSEE	1x	10x	0.5ml	1ml	1.5ml	2ml
PMSF[a]	0.5mM	200mM @4°C	12.5µl	25µl	37.5µl	50µl
Aprotinin	10µg/ml	10µg/µl @4°C	5µl	10µl	15µl	20µl
Leupeptin	10µg/ml	10µg/µl @-20°C	5µl	10µl	15µl	20µl
Pepstatin A	2µg/ml	1µg/µl @-20°C	10µl	20µl	30µl	40µl
Na_3VO_4	1mM	200mM @-20°C	25µl	50µl	75µl	100µl
NaF	1mM	200mM @RT	25µl	50µl	75µl	100µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

10x XIP-TSEE*			
Component	Final Concentration	Stock	50 ml
Water			23.375ml
Tris•Cl (pH7.5)	200mM	1M, pH7.5 @RT	10ml
NaCl	1.5M	5M @RT	15ml
EDTA (pH8.5)	10mM	0.5M @4°C	1ml

EGTA (pH8.0)	5mM	0.4M @4°C	625µl

* This solution can be stored temporary at @4°C, or long term @-20°C

XIP-SonicBuffer*						
Component	Final Concentration	Stock	1ml	3ml	5ml	10ml
Water			483.5µl	1450.5µl	2.425ml	4.84ml
2x XIP-SonicBuf-GTDSTSEE	1x	2x	500µl	1500µl	2.5 ml	5ml
PMSF[a]	0.5mM	200mM @4°C	2.5µl	7.5µl	12.5µl	25µl
Aprotinin	10µg/ml	10µg/µl @4°C	1µl	3µl	5µl	10µl
Leupeptin	10µg/ml	10µg/µl @-20°C	1µl	3µl	5µl	10µl
Pepstatin A	2µg/ml	1µg/µl @-20°C	2µl	6µl	10µl	20µl
Na_3VO_4	1mM	200mM @-20°C	5µl	15µl	25µl	50µl
NaF	1mM	200mM @RT	5µl	15µl	25µl	50µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

2x XIP-SonicBuf-GTDSTSEE*			
Component	Final Concentration	Stock Solution	30ml
Water			9ml
Glycerol	20%	100%	6ml
TritonX-100	2%	10%	6ml
DOC	0.8%	10%	2.4ml
SDS	0.2%	10%	600µl
10x TSEE	2x	10x	6ml

* This solution can be stored temporary at @4°C, or long term @-20°C

Bioruptor Sonication buffer*				
Comp	Fin. Con	Stock	1ml	10ml
Water			847.5µl	8.48ml
HEPES-Na (pH 7.5)	20mM	1M	20µl	200µl
EDTA (pH 8.0)	2mM	500mM	4µl	40µl
SDS	1%	10%	100µl	1000µl
Triton X-100	0.2%	10%	20µl	200µl
Spermidine	0.5mM	0.5M	1.0µl	10µl
Spermine	0.15mM	0.15M	1.0µl	10µl
PMSF[a]	0.5mM	200mM @4°C	2.5 µl	25µl

Aprotinin	10µg/ml	10µg/µl @4°C	1.0µl	10µl
Leupeptin	10µg/ml	10µg/µl @-20°C	1.0µl	10µl
Pepstatin A	2µg/ml	1µg/µl @-20°C	2.0µl	20µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

XIP-IPBuffer*					
Comp	Fin. Con	Stock	10ml	20ml	30ml
Water			4.84ml	9.68ml	14.52ml
2x XIP-IPBuf-TSTSEE	1x	2x	5ml	10ml	15ml
PMSF[a]	0.5mM	200mM @4°C	25µl	50µl	75µl
Aprotinin	10µg/ml	10µg/µl @4°C	10µl	20µl	30µl
Leupeptin	10µg/ml	10µg/µl @-20°C	10µl	20µl	30µl
Pepstatin A	2µg/ml	1.0µg/µl @-20°C	20µl	40µl	60µl
Na_3VO_4	1mM	200mM @-20°C	50µl	100µl	150µl
NaF	1mM	200mM@RT	50µl	100µl	150µl

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

Blocking buffer				
Component	Final Concentration	Stock Solution	1ml	10ml

XIP-IPBuffer			850 µl	8.5ml
Sonicated Salmon sperm DNA*	0.1µg/µl	1µg/µl	100µl	1ml
BSA	5µg/µl	100µg/µl	50µl	500µl

* For ChIP-seq add 0.1µg/µl yeast tRNA instead

2x XIP-IPBuf-TSTSEE*				
Component	Final Concentration	Stock Solution	30ml	50ml
Water			17.94ml	29.9ml
1) TritinX-100	2%	10%	6.0ml	10ml
2) SDS	0.02%	10%	60µl	100µl
3) 10x TSEE	2x	10x	6.0ml	10ml

* This solution can be stored temporary at @4°C, or long term @-20°C

XIP-HiSalt Buffer*					
Component	Final Concentration	Stock Solution	10ml	20ml	30ml
Water			3.835ml	7.67ml	11.505ml
2x XIP-HiLoSalt-TTEE	1x	2x	5ml	10ml	15ml
PMSF[a]	0.5mM	200mM @4°C	25µl	50µl	75µl
Aprotinin	10µg/ml	10µg/µl @4°C	10µl	20µl	30µl

Leupeptin	10µg/ml	10µg/µl @-20°C	10µl	20µl	30µl
Pepstatin A	2µg/ml	1µg/µl @-20°C	20µl	40µl	60µl
Na_3VO_4	1mM	200mM @-20°C	50µl	100µl	150µl
NaF	1mM	200mM @RT	50µl	100µl	150µl
NaCl	500mM	5M	1.0ml	2.0ml	3.0ml

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

XIP-LowSalt Buffer*				
Component	Final Concentration	Stock Solution	10ml	20ml
Water			3.835ml	7.67ml
2x XIP-HiLoSalt-TTEE	1x	2x	5ml	10ml
PMSF[a]	0.5mM	200mM @4°C	25µl	50µl
Aprotinin	10µg/ml	10µg/µl @4°C	10µl	20µl
Leupeptin	10µg/ml	10µg/µl @-20°C	10µl	20µl
Pepstatin A	2µg/ml	1µg/µl @-20°C	20µl	40µl
Na_3VO_4	1mM	200mM@-20°C	50µl	100µl
NaF	1mM	200mM@RT	50µl	100µl
Water			1.0ml	2.0ml

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

2x XIP-HiLoSalt-TTEE*			
Component	Final Concentration	Stock Solution	50ml
Water			37.68ml
1) Tris•Cl pH7.5	40mM	1M pH7.5 @RT	2.0ml
2) TritinX-100	2%	10%	10ml
3) EDTA pH8.5	2mM	0.5M @4°C	200µl
4) EGTA pH8.0	1mM	0.4M @4°C	125µl

* This solution can be stored temporary at @4°C, or long term @-20°C

Li Buffer*				
Component	Final Concentration	Stock Solution	10ml	20ml
Water			7.4 ml	14.98 ml
Tris•Cl pH7.5	10mM	1M pH7.5 @RT	100 µl	200 µl
Igpel CA-630	1%	10%	1.0ml	2.0ml
DOC[a]	1%	10%	1.0ml	2.0ml
LiCl	250mM	8M	312.5µl	625µl
EDTA	1mM	500mM	20µl	40µl
PMSF[b]	0.5mM	200mM @4°C	25µl	50µl
Aprotinin	10µg/ml	10µg/µl @4°C	10µl	20µl
Leupeptin	10µg/ml	10µg/µl @-20°C	10µl	20µl

Pepstatin A	2µg/ml	1µg/µl @-20°C	20µl	40µl
Na_3VO_4	1mM	200mM @-20°C	50µl	100µl
NaF	1mM	200mM @RT	50µl	100µl

*Make fresh each time

[a]Warm up 10% DOC at 25°C before use since it will precipitate @ RT

[b]PMSF is not stable in aqueous solution, add PMSF to the solution just before use

Elution buffer*				
Component	Final Concentration	Stock Solution	1ml	2ml
Water			0.8ml	1.6ml
SDS	1%	10%	0.1ml	0.2ml
$NaHCO_3$	100mM	1.0M*	0.1ml	0.2ml

*Make fresh each time

[a]PMSF is not stable in aqueous solution, add PMSF to the solution just before use